

AD-A198 477

INVITATION TO THE SYMPOSIUM

Murray A. Wornat

APPLIED RESEARCH LABORATORIES  
THE UNIVERSITY OF TEXAS AT AUSTIN  
POST OFFICE BOX 8029, AUSTIN, TEXAS 78713-8029

8 February 1988

Technical Report

Approved for public release;  
distribution is unlimited.

**BEST  
AVAILABLE COPY**

Prepared for:

APPLIED RESEARCH LABORATORIES  
THE UNIVERSITY OF TEXAS AT AUSTIN  
POST OFFICE BOX 8029, AUSTIN, TX 78713-8029



**DTIC  
ELECTE**  
AUG 31 1988  
**S D**  
**E**

88 8 31 085

# UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

## REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 0704-0188

1a REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>			1b RESTRICTIVE MARKINGS		
2a SECURITY CLASSIFICATION AUTHORITY			3 DISTRIBUTION / AVAILABILITY OF REPORT  Approved for public release; distribution is unlimited.		
2b DECLASSIFICATION / DOWNGRADING SCHEDULE			5 MONITORING ORGANIZATION REPORT NUMBER(S)		
4 PERFORMING ORGANIZATION REPORT NUMBER(S)  ARL-TR-88-7			7a NAME OF MONITORING ORGANIZATION  Applied Research Laboratories		
6a NAME OF PERFORMING ORGANIZATION  Applied Research Laboratories		6b OFFICE SYMBOL (If applicable)  ARL:UT	7b ADDRESS (City, State, and ZIP Code) The University of Texas at Austin P.O. Box 8029 Austin, Texas 78713-8029		
8a NAME OF FUNDING / SPONSORING ORGANIZATION		8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER ARL:UT Independent Research and Development Program		
8c ADDRESS (City, State, and ZIP Code)		10 SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
11 TITLE (Include Security Classification)  Invitation to the Bispectrum					
12 PERSONAL AUTHOR(S) Wolinsky, Murray A.					
13a TYPE OF REPORT technical		13b TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Year, Month, Day) 88-2-8	
15 PAGE COUNT 84					
16 SUPPLEMENTARY NOTATION					
17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP			
			bispectrum ; higher-order spectrum		
			cross-spectrum ; non-Gaussianity		
			frequency coupling ; nonlinear signals		
19 ABSTRACT (Continue on reverse if necessary and identify by block number)					
<p>The bispectrum is examined from a variety of physical and mathematical perspectives. Particular stress is placed on those examples or approaches which aid intuitive understanding of the bispectrum and its properties. A key contribution of this work is the discovery and detailed development of a time series model which can reproduce any possible bispectrum. This model is used to discuss practical concerns relating to the bispectrum such as sampling and estimation issues.</p>					
20 DISTRIBUTION / AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS			21 ABSTRACT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>		
22a NAME OF RESPONSIBLE INDIVIDUAL Marshall E. Frazer			22b TELEPHONE (Include Area Code) 512-835-3450		22c OFFICE SYMBOL SPG

**UNCLASSIFIED**

<b>APPENDIX A</b>	<b>Catalog of Examples</b>	<b>73</b>
A-1	Linear Non-Gaussian White Noise Passed through Linear Device .	74
A-2	Gaussian Noise Passed through Squarer. . . . .	74
A-3	Poisson Activity . . . . .	74
A-4	Interacting Sinusoids . . . . .	75
A-5	Generalization of A-4, due to Powers. . . . .	75
A-6	A Realistic Evolution Equation, due to Powers. . . . .	77
A-7	A Stochastic Model of Gear Noises . . . . .	78
A-8	Brillinger's Polyspectrum Example. . . . .	78
<b>REFERENCES</b>		<b>80</b>

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



## LIST OF FIGURES

1-1	Localization Via Cross-spectrum. . . . .	4
1-2	Localization Via Bispectrum (Munk's Version). . . . .	5
1-3	Localization Via Bispectrum (MacDonald's Version). . . . .	6
2-1	Fundamental Domain of the Bispectrum. . . . .	14
2-2	$B(\omega_1, \omega_2) = B(\omega_2, \omega_1)$ . . . . .	16
2-3	$B(\omega_1, \omega_2) = B(-\omega_1 - \omega_2, \omega_2)$ . . . . .	16
2-4	$B(\omega_1, \omega_2) = B(2\pi - \omega_1 - \omega_2, \omega_1)$ . . . . .	18
2-5	$B(\omega_1, \omega_2) = B(2\pi - \omega_1 - \omega_2, \omega_2)$ . . . . .	19
2-6	$B(\omega_1, \omega_2) = [B(-\omega_1, -\omega_2)]^*$ . . . . .	19
3-1	The $c_3$ Plane in Terms of the Weights $g$ . . . . .	27
3-2	Integration Limits before Change of Variables. . . . .	33
3-3	Integration Limits after Change of Variables. . . . .	34
3-4	Migrating Triangular Piece (before Migration). . . . .	35
3-5	Migrating Triangular Piece (after Migration). . . . .	36
3-6	Commutative Diagram for Universal Bispectrum Model. . . . .	41
5-1	Artificial Bispectrum Example. . . . .	59
5-2	The "Predicted" Bicorrelations. (Test Model) . . . . .	61
5-3	The "Observed" Bicorrelations. (Test Model) . . . . .	62
5-4	Predicted and Observed Correlations. (Test Model) . . . . .	63
5-5	Predicted and Observed Power Spectra. (Test Model) . . . . .	64
5-6	Predicted and Observed Bispectra. (Test Model) . . . . .	65
5-7	Predicted and Observed Bispectra. (Square Model) . . . . .	66
5-8	Predicted and Observed Power Spectra. (Square Model) . . . . .	67
5-9	Predicted and Observed Bicorrelations. (Square Model) . . . . .	68
A-1	Interacting Sinusoids. . . . .	76

## TABLE OF CONTENTS

<b>LIST OF FIGURES</b>	vii
<b>1 The Bispectrum in Multiple Perspectives</b>	<b>1</b>
1-1 Intuitive Understanding of the Bispectrum is Possible. . . . .	1
1-2 Bispectrum = Double Fourier Transform of Third Order Cumulant (Definition). . . . .	2
1-3 Bispectrum = Expectation of Product of Interacting Frequency Components. . . . .	3
1-4 Bispectrum = Cross-Spectrum of Frequency Components (Mac- Donald's Filter Interpretation). . . . .	4
1-5 Bispectrum = Nonlinear Coupling of Product of Interacting Fre- quency Components. . . . .	6
1-6 Bispectrum = Rate of Transfer of Energy. . . . .	9
1-7 Bispectrum = Decomposition of Cube of Time Series. . . . .	11
1-8 The Bispectrum Can Be Interpreted in a Variety of Ways. . . . .	12
<b>2 Mathematical Properties of the Bispectrum</b>	<b>13</b>
2-1 The Principal Domain of a Bispectrum is a Consequence of Sta- tionarity, "Reality" , and Symmetries. . . . .	13
2-2 For a Properly Sampled Continuous Time Signal, the "Extra" Tri- angle is "Forbidden". . . . .	17
2-3 The Most Basic Mathematical Properties of the Bispectrum are those which Determine the Fundamental Domain. . . . .	21
<b>3 The Universal Bispectrum Model</b>	<b>22</b>
3-1 The Universal Bispectrum Model is the Simplest Model that can Reproduce any Possible Bispectrum. . . . .	22
3-2 The Second-order Cumulant for the Universal Bispectrum Model is Rather Messy and Not Very Informative. . . . .	23
3-3 For the Universal Bispectrum Model, the Third-Order Cumulant Equals $\sigma^4$ times the Weight. . . . .	24

3-4	The Power Spectrum of the Universal Bispectrum Model is Somewhat Informative. . . . .	26
3-5	The Bispectrum of the Universal Bispectrum Model is Given by the Frequency Coupling Coefficients. . . . .	28
3-6	Inversion of the Model is Easy. . . . .	29
3-7	The Universal Bispectrum Model in the Frequency Domain is as Simple as it Can Possibly Be. . . . .	32
3-8	The Bispectrum Can Be Computed Directly in the Frequency Domain. . . . .	36
3-9	The Universal Bispectrum Model Might Arise in Practice, Almost. . . . .	39
3-10	The Universal Bispectrum Model Can Be Clarified by a Diagram. . . . .	40
<b>4</b>	<b>Estimation of the Bispectrum</b>	<b>42</b>
4-1	At Least Four Techniques are Used To Estimate the Bispectrum. . . . .	42
4-2	One Must Distinguish the Sample Fourier Transform from the Quantity it Estimates - the Population Fourier Transform. . . . .	43
4-3	Fanfare for Brillinger's Formula. . . . .	44
4-4	A Diversion: Cumulants are Simply Expectations with Lower-Order Dependence Removed. . . . .	46
4-5	Everything We Need To Know: Table of Zero Mean Cumulants . . . . .	48
4-6	The Practice: Previous Results Allow One To Estimate the Power Spectrum. . . . .	48
4-7	The Payoff: Previous Results Allow One To Estimate the Bispectrum. . . . .	51
4-8	The Mathematical Results Can Be Turned into a Practical Procedure for Estimating the Bispectrum. . . . .	53
4-9	The Practical Procedure Can Be Simplified To Give Order-of-Magnitude Estimates. . . . .	56
4-10	Estimating the Bispectrum Requires Smoothing. . . . .	56
<b>5</b>	<b>The Model in Practice</b>	<b>58</b>
5-1	A Simple Example is Given by Choosing the Bispectrum Constant in a Specified Rectangular Region and Zero Elsewhere. . . . .	58
5-2	Computer Simulations Give Good Agreement with the Model. . . . .	60
5-3	Reasonable Parameter Values are Obtainable through a Combination of Mathematical, Statistical, and Practical Considerations. . . . .	67
5-4	The Model Works! . . . . .	70
<b>6</b>	<b>Summary</b>	<b>72</b>

## CHAPTER 1

### The Bispectrum in Multiple Perspectives

#### 1-1 Intuitive Understanding of the Bispectrum is Possible.

The bispectrum is a potentially valuable tool for analyzing any nonlinear process. Moreover, it has been applied to a wide variety of nonlinear problems in numerous disciplines. [1.1-1] However, only recently has it gained a substantial measure of popularity.

There are at least four possible reasons as to why the bispectrum is only now coming into its own. First, the bispectrum is computationally relatively expensive. Second, the additional knowledge which the bispectrum supplies may be too limited to justify the extra effort required. Third, physical and intuitive understanding of the bispectrum are not widespread. Fourth, the mathematics of the bispectrum are sometimes considered excessively complicated.

The computational expense of the bispectrum is no longer a significant problem as powerful machines have become commonplace. The impact of the second reason is rather difficult to assess: it is true that many investigations do not require the additional insight afforded by the bispectrum and further, in many cases, this additional information is small. However, it is not easy to know in advance that the information gained will or will not be of value. Moreover, in certain circumstances *any* information may be of such importance that the possibility of its acquisition should not be ruled out at the outset. Finally, the literature of the applied bispectrum demonstrates its usefulness — at least to the specific problems studied.

This report focuses on the third and fourth impediments to the more widespread use of the bispectrum: lack of physical, intuitive and mathematical understanding of the bispectrum.

The remainder of this chapter will attempt to foster the reader's intuitive and physical understanding of the bispectrum by presenting the bispectrum in different perspectives. The hope is that of the following assortment of ways of

regarding the bispectrum, at least one will "make sense" to the reader. More optimistically, these viewpoints will reinforce one another to provide a satisfactory intuitive picture. These sections do not contain detailed mathematics and the reader is encouraged to avoid any perspective which seems unnatural or unduly complicated. For this chapter, no results are proved and the mathematics are partially *ornamental* (and should be studied only if helpful). (Proofs can be found in the references, if desired.) However, more mathematical detail is available in the following chapters.

The basic mathematics of the bispectrum is described in general in Chapter 2 and a specific model designed to reinforce this mathematical understanding is presented in Chapter 3. Practical considerations such as how to estimate the bispectrum of a time series are dealt with in Chapter 4. (Part of the purpose of this report is to make the bispectrum a more useful tool, so such considerations need to be included.) The fifth chapter presents a detailed worked-out example and uses the model presented in Chapter 3 and the statistical methods of Chapter 4. The final chapter presents a summary of the report and an appendix presents a catalog of time series with corresponding bispectra.

The bispectrum comes in many varieties — deterministic or stochastic, discrete or continuous time, and with or without the assumption of stationarity. The bulk of the literature considers the stationary stochastic case and this report shall dwell primarily on this case as well. Moreover, time shall in general be taken to be continuous.

## 1-2 Bispectrum = Double Fourier Transform of Third Order Cumulant (Definition).

The power spectrum  $P(\omega)$  of a continuous time, stationary stochastic time series  $x(t)$  is usually defined as the Fourier transform of the second-order covariance function  $c_2(\tau)$ . Explicitly one has

$$c_2(\tau) = \langle x(t) x(t + \tau) \rangle \quad (1-2.1)$$

and

$$P(\omega) = \int_{-\infty}^{\infty} c_2(\tau) e^{-i\omega\tau} d\tau, \quad (1-2.2)$$

where the angle brackets denote ensemble average and this average is independent of the time  $t$  because of the assumed stationarity.

The bispectrum  $B(\omega_1, \omega_2)$  is defined by analogy to the power spectrum [1.2-1]. It is the (double) Fourier transform of the third-order covariance function  $c_3(\tau_1, \tau_2)$ .



Explicitly,

$$c_3(\tau_1, \tau_2) = \langle x(t) x(t + \tau_1) x(t + \tau_2) \rangle, \quad \text{and} \quad (1-2.3)$$

$$B(\omega_1, \omega_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c_3(\tau_1, \tau_2) e^{-i(\omega_1 \tau_1 + \omega_2 \tau_2)} d\tau_1 d\tau_2. \quad (1-2.4)$$

What kind of insight does the definition provide into the interpretation of the bispectrum? Unfortunately, *precious little* seems to be the answer. Features in the bispectrum correspond to effects of (triple) correlation in the original signal. It is clear that the bispectrum is zero if there is no triple correlation. Using one's intuition about the relationship between time and frequency domain quantities, it is clear that a "spiky" bispectrum implies a rather "flat" triple correlation and therefore a time series possessing intervals of slow change. Inversely, a repetitious series of spikes in the time domain may be expected to correspond to flatter peaks in the bispectrum. Beyond this much, the definition seems not to provide aid to the intuition.

### 1-3 Bispectrum = Expectation of Product of Interacting Frequency Components.

If one transforms the original signal  $x(t)$  into the frequency domain by defining  $\hat{x}(\omega)$  such that

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{x}(\omega) e^{i\omega t} d\omega, \quad (1-3.1)$$

then it is well-known (and easy to prove) that the power spectrum can be written directly in terms of the frequency domain quantities

$$P(\omega) = \langle \hat{x}(\omega) \hat{x}(-\omega) \rangle. \quad (1-3.2)$$

Again the angle brackets denote ensemble average. The fact that the two frequencies  $\omega$  and  $-\omega$  sum to zero is a direct consequence of stationarity. It is easy to extend the proof of the above formula to apply to the bispectrum.

One finds that

$$\boxed{B(\omega_1, \omega_2) = \langle \hat{x}(\omega_1) \hat{x}(\omega_2) \hat{x}(-\omega_1 - \omega_2) \rangle.} \quad (1-3.3)$$

The fact that the three frequencies  $\omega_1$ ,  $\omega_2$ , and  $-\omega_1 - \omega_2$  sum to zero is a direct consequence of stationarity.

This form of the bispectrum is somewhat helpful to the intuition. The bispectrum can only be nonzero at a particular pair of frequencies  $(\omega_1, \omega_2)$  if the

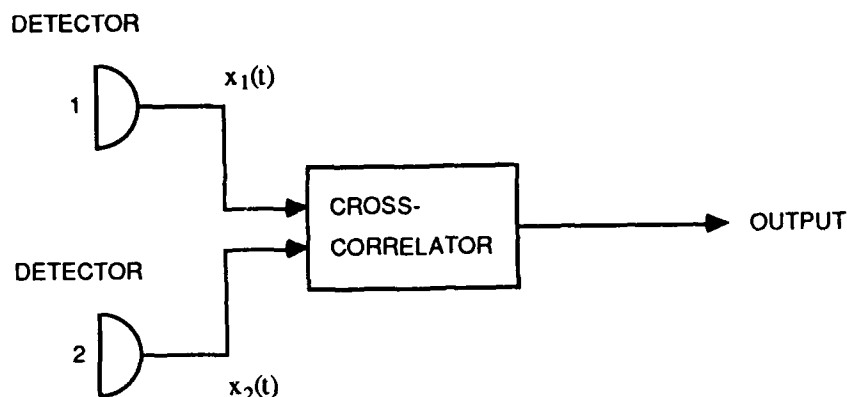


Figure 1-1 LOCALIZATION VIA CROSS-SPECTRUM.

frequency component at the sum of these frequencies ( $\hat{x}(\omega_1 + \omega_2)$ ) is statistically dependent on the product ( $\hat{x}(\omega_1) \hat{x}(\omega_2)$ ) of these frequency components. That is, the two original frequencies must *couple* through their sum term. Moreover this is the only (*three-wave*) coupling possible for stationary time series. (The previous statement is true provided the time series is continuous-time, as has been assumed so far. For a discrete time series it is only necessary for the three frequencies to sum to zero modulo  $2\pi$ .)

Thus the bispectrum measures coupling between frequencies or, equivalently, it measures a specific phase relation between frequency components. If there was no such relationship between the phases then the expectation above would vanish.

#### 1-4 Bispectrum = Cross-Spectrum of Frequency Components (MacDonald's Filter Interpretation).

In an interesting MITRE company report [1.4-2] devoted to the bispectrum, Munk discusses the possibility of using the bispectrum to localize a target. This possibility arises because of the following interesting analogy.

Currently cross-spectral analysis is frequently used for localization, as shown in Figure 1-1. One has two detectors and consequently two signals. These signals can be crosscorrelated and analyzed to determine phase differences. (The cross-spectrum can be computed by either of two methods. One can compute the crosscorrelation first and then Fourier transform or one can compute the transforms first and then multiply.) In bispectral analysis, there is only one

signal at the outset, but it is separated into two frequency components. Munk suggests that one can then heterodyne both components to zero frequency and finally crosscorrelate. (See Figure 1-2.)

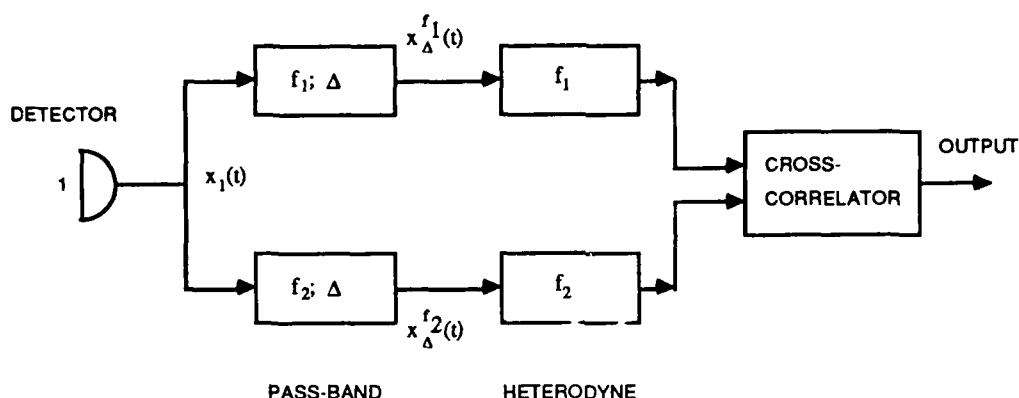
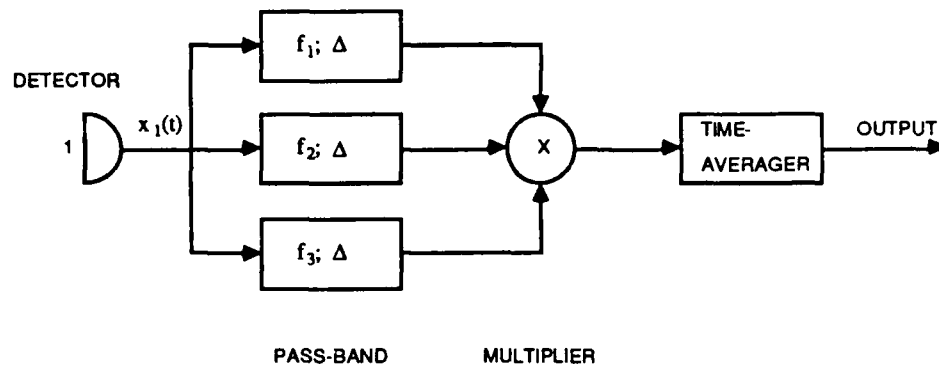


Figure 1-2 LOCALIZATION VIA BISPECTRUM (MUNK'S VERSION).

Here one can see that bispectral analysis looks at just that information which ordinary spectral analysis throws away: inter-frequency phase information. (The "relative phase" or "inter-frequency phase" of the two frequencies  $f_1$  and  $f_2$  is generally defined as the phase part of the bispectrum.)

MacDonald [1.4-3] provides a formal treatment of the above argument. A condensed version of his argument is presented below and may be skipped if desired. MacDonald's treatment actually leads to the representation shown in Figure 1-3 which is slightly different from Munk's version. In fact, there is no explicit treatment of the bispectrum which corresponds exactly to Munk's description and it would be interesting to try to construct one. (This might make a valuable mini-research project for the reader.)

Start with a continuous time, stationary, ergodic, zero mean, stochastic time series  $x(t)$ . Imagine that one can construct ideal passband filters with center frequency  $f_1$  and bandwidth  $\Delta$ . If  $x(t)$  is input to the passband filter, denote the filter output by  $x_{\Delta}^{f_1}(t)$ . Now take the input time series  $x(t)$  and "fan out" into three copies. Send one copy into a passband filter centered at  $f_1$ . Send another copy into a passband filter centered at  $f_2$ . Send the final copy into a passband filter centered at  $f_3 = f_1 + f_2$ . Let all three filters have bandwidth  $\Delta$ . Now multiply the outputs of all three filters together. In a sense, the fact that the third frequency is the sum of the other two allows one to think of this multiplication as heterodyning both signals to zero and crosscorrelating in the same step. (This is the part that



**Figure 1-3** LOCALIZATION VIA BISPECTRUM (MACDONALD'S VERSION).

needs work to give a completely explicit meaning to Munk's interpretation.) Take the output of the multiplier and time-average. Finally, multiply the output of the "time-average" by  $2/(3\Delta^2)$ . The result is the real part of the bispectrum of  $x(t)$  evaluated at  $f_1, f_2$ . If the imaginary part is desired, it is only necessary to change the  $f_3$  passband filter to include a uniform  $\pi/2$  phase shift.

This condensed version of MacDonald's treatment may be viewed as a signal processing or *analog* explanation of the bispectrum (in contrast to the digital explanation of 1.3).

#### 1-5 Bispectrum = Nonlinear Coupling of Product of Interacting Frequency Components.

This section shall relate the bispectrum to two models involving nonlinear coupling of waves. The first model consists of a simple quadratic nonlinearity. The second involves a physically realistic nonlinear evolution equation. Both models are presented by E. J. Powers and his co-authors [1.5-1, 1.5-2].

Establishing a suitable framework for developing both applications is first on the agenda. Some noisy quantity of interest (call it  $\phi$  and for simplicity take it to be scalar) propagates in some uniform medium. For simplicity of exposition  $\phi$  will be assumed to vary significantly in only one dimension, e.g., the  $x$  direction. Following the usual procedure, Fourier decomposition of  $\phi$  is called for. Write

$$\phi(x, t) = \sum_{\omega} \phi_{\omega}(x) \exp[ik(\omega)x - i\omega t]. \quad (1-5.1)$$

The quantity  $k(\omega)$  is the frequency dependent wave number and characterizes the medium under consideration. Since  $\phi$  is a random quantity, so is the  $\phi_{\omega}(x)$ .

Interest centers in both models on the behavior in time of the values of  $\phi$  at one particular point  $x_p$  in space, the *probe* position. To expedite matters, define  $y(t)$  to equal  $\phi(x_p, t)$ . If one sets

$$Y(\omega) \equiv \phi_\omega \exp[ik(\omega)x_p], \quad (1-5.2)$$

then  $Y$  and  $y$  form a Fourier transform pair.

The power spectrum of  $y$ , as usual, is

$$P(\omega) = \langle Y(\omega) Y^*(\omega) \rangle \quad (1-5.3)$$

and the bispectrum is

$$B(\omega_1, \omega_2) = \langle Y(\omega_1) Y(\omega_2) Y^*(\omega_1 + \omega_2) \rangle. \quad (1-5.4)$$

Both models will be simplified in that only three waves will be considered. These waves will be assigned frequencies  $\omega$ ,  $\omega_1$ , and  $\omega_2$ , where  $\omega = \omega_1 + \omega_2$ .

With the framework for both models established, one can now consider the first model [1.5-1]. In this model, the nonlinearity is taken to be of the simple form

$$Y(\omega) = AY(\omega_1)Y(\omega_2) + Y'. \quad (1-5.5)$$

Here  $Y'$  is some quantity which contributes to  $Y(\omega)$  but is statistically independent of the coupling term.  $A$  represents the coupling constant.

Computing the power spectrum at  $\omega$ , one easily<sup>1</sup> finds

$$P(\omega) = |A|^2 \langle |Y(\omega_1) Y(\omega_2)|^2 \rangle + \langle |Y'|^2 \rangle. \quad (1-5.6)$$

The first term represents the contribution of the wave-wave coupling interaction and the second includes everything else.

It is just as easy<sup>2</sup> to compute the bispectrum at  $\omega_1, \omega_2$ :

$$B(\omega_1, \omega_2) = A^* \langle |Y(\omega_1) Y(\omega_2)|^2 \rangle. \quad (1-5.7)$$

One sees immediately that the bispectrum conveys information only about the coupling interaction. In fact, one could partially normalize the bispectrum so as to get the coupling coefficient from experimental data as follows:

$$A = \frac{B^*(\omega_1, \omega_2)}{\langle |Y(\omega_1) Y(\omega_2)|^2 \rangle}. \quad (1-5.8)$$

<sup>1</sup>Proving this takes only a few lines and may be a useful exercise.

<sup>2</sup>This formula is also very easy to derive and might serve well as the reader's first "bispectral computation".

In the second model [1.5-2], a realistic evolution equation is assumed,

$$\frac{\partial \phi_\omega}{\partial x} = V \phi_{\omega_1} \phi_{\omega_2} \exp(i\delta_k x), \quad (1-5.9)$$

with frequency matching ( $\omega = \omega_1 + \omega_2$ ), possible mismatch in wave number (characterized by  $\delta_k \equiv k_{\omega_1} + k_{\omega_2} - k_\omega$ ), and wave-wave coupling (with coupling coefficient  $V$ ).

This type of equation is commonly found in nonlinear optics and in plasma physics. It shows how a wave gains or loses energy due to interactions with waves of different frequencies as it propagates in space.

As implied earlier, one should properly

- make the right side an integral over all  $\omega_1$  (taking  $\omega_2 = \omega - \omega_1$ ),
- let  $V$  be dependent on  $\omega$ ,  $\omega_1$ , and  $\omega_2$ , and
- take this equation to be true for all  $\omega$ .

It will be easier (as above) to interpret the above equation as true for some fixed  $\omega$  and to assume that for the wave at this frequency  $\omega$  there is one pair of frequencies ( $\omega_1$  and  $\omega_2$ ) that dominates the integral. In particular the frequency dependence of  $V$  can be omitted.

In terms of the quantities of experimental interest, the above equation becomes

$$\frac{\partial Y_\omega}{\partial x} = V Y_{\omega_1} Y_{\omega_2} + ik(\omega)Y(\omega). \quad (1-5.10)$$

It is a simple exercise (left to the reader) to take the above equation and derive

$$\frac{\partial P(\omega)}{\partial x} = V B(\omega_1, \omega_2) + V^* B^*(\omega_1, \omega_2). \quad (1-5.11)$$

One should now follow Powers and his co-authors and re-write the above equation in terms of strictly real quantities  $V_R$ ,  $V_I$ ,  $B_R$ , and  $B_I$ , where

$$2V \equiv V_R + iV_I, \quad \text{and} \quad B \equiv B_R + iB_I. \quad (1-5.12)$$

This gives the wonderful equation

$$\boxed{\frac{\partial P(\omega)}{\partial x} = V_R B_R(\omega_1, \omega_2) + V_I B_I(\omega_1, \omega_2).} \quad (1-5.13)$$

This equation demonstrates that power is put into the wave at frequency  $\omega$  (as one advances in the  $x$  direction) due to coupling with the waves at  $\omega_1$  and  $\omega_2$ . Further, the real part of the bispectrum gives the amount of that coupling due to the real part of the coupling coefficient  $V$  (and similarly for the imaginary part).

## 1-6 Bispectrum = Rate of Transfer of Energy.

Since the bispectrum measures the amount of coupling between two frequencies and since this coupling governs the rate at which energy is exchanged between these frequencies, it follows that the bispectrum provides information regarding how energy moves from one frequency range to another. This has been demonstrated in the previous section. This information is of particular interest in at least two disciplines: the study of turbulence and the study of nuclear fusion in plasmas.

This section shall present a brief exposition of the theory of homogeneous turbulence. Much of the experimental bispectral literature deals with this problem because the bispectrum is so intimately related to the quantities of experimental and theoretical interest. If one did not know better, one might even suspect that this entire subject was invented to provide an application for bispectra. In a sense, the discussion in this section is the *duai* of the previous section in that here one considers energy transfer in *time* rather than in *space* and that here the bispectrum is extended to functions of *space* rather than *time*. (The reader should feel free to *skim* the remainder of this section. He/she is encouraged *not* to *skip* it entirely.)

In homogeneous turbulence, one is concerned with an infinite body of fluid and its random motions [1.6-1]. These motions are taken to be independent of position on average, and hence *homogeneous*. The fluid is taken to be incompressible and can be characterized by its (uniform, i.e., position-independent) density  $\rho$  and its (uniform) kinematic viscosity  $\nu$ . This is of course a vast simplification of typical situations. However, even this simple problem evades mathematical solution. Moreover, one can experimentally produce a good approximation to homogeneous turbulence by passing a fluid perpendicularly through an array of metal bars ("grid-generated turbulence") [1.6-1, 1.6-2].

The state of the fluid is specified by providing its velocity  $\vec{v}$  and pressure  $p$  as functions of time  $t$  and space  $\vec{x}$ . The fluid motion is assumed to be random so that one is primarily concerned with expectations and especially correlations in time and/or in space. For example, one may desire  $\langle v_x(\vec{x}_1, t) v_x(\vec{x}_2, t) v_x(\vec{x}_3, t) \rangle$  or  $\langle v_x(\vec{x}, t_1) v_x(\vec{x}, t_2) \rangle$ .

The motion of the fluid is governed by the Navier-Stokes equation,

$$D\vec{v} = -\frac{1}{\rho}\nabla p + \nu\nabla^2\vec{v}, \quad (1-6.1)$$

and is subject to the incompressibility constraint

$$\nabla \cdot \vec{v} = 0. \quad (1-6.2)$$

The operator  $D$  is the *kinematic derivative* and is defined as

$$D \equiv \frac{\partial}{\partial t} + \vec{v} \cdot \nabla. \quad (1-6.3)$$

The requirement of homogeneity (averages are independent of position) supplies the boundary conditions (in an unusual, though pleasant, fashion).

Turbulence is inherently a "multi-scale" phenomenon. It is caused by interaction of phenomena which take place on very different length scales (from clearly visible to microscopic). Thus a decomposition of the motion of the fluid velocity in terms of motions at different length scales is suggested. One is led to write

$$\vec{v}(\vec{x}, t) = \int \exp(i\vec{k} \cdot \vec{x}) d\vec{z}(\vec{k}, t). \quad (1-6.4)$$

This is a Cramér or Fourier-Stieltjes representation of the velocity field. The mathematics (which will be dealt with in Chapter 3) is not important here: all that is required is the understanding that the velocity is written in terms of motions which are on varied length scales. The vector  $\vec{k}$  specifies the scale of the motion – the greater the magnitude of  $\vec{k}$  the smaller the scale (wavelength). The integral is taken over all possible scales and is evaluated at a particular time. This makes it possible to ask questions such as what portion of the total energy of the fluid motion is due to motions on a particular length scale and how does this fraction evolve with time. In other words, one can trace the transfer of energy in time from one spatial frequency range to another.

It must be stressed that the interest is in *spatial* frequencies rather than temporal ones. One is concerned primarily with velocity as a function of position at a particular time rather than with velocity at a particular place as a function of time. However, this fact makes no difference to the mathematics. In particular, the homogeneity assumption is the spatial analog of the requirement of stationarity, so the "space" series have the usual properties of time series. Fourier transforming the Navier-Stokes equation and simplifying gives an equation for the dynamics of spectral energy transfer,

$$\frac{\partial}{\partial t} \left[ \frac{1}{2} \Phi_{nn}(\vec{k}, t) \right] = \int T(\vec{k}, \vec{k}', t) d\vec{k}' - \nu k^2 \Phi_{ii}(\vec{k}, t). \quad (1-6.5)$$

In this equation,  $\Phi_{nn}(\vec{k}, t)$  is the spectral energy density (at a particular time),  $T$  is the energy transfer function, and  $k$  is the magnitude of  $\vec{k}$ . If  $T$  is known then one can say how energy in a particular (spatial) frequency range changes in time due to transfer from motions at other (spatial) frequencies.



The spectral energy density  $\Phi$  is defined in terms of the velocity components as

$$\Phi_{ij}(\vec{k}, t) dk_1 dk_2 dk_3 = \langle d\tilde{z}_i^*(\vec{k}, t) d\tilde{z}_j(\vec{k}, t) \rangle. \quad (1-6.6)$$

As one might suspect from the introduction to this section, the energy transfer function  $T$  is simply expressed in terms of the (spatial) bispectrum<sup>3</sup> of the velocity-field:

$$T(\vec{k}, \vec{k}', t) = -\vec{k} \cdot \text{Im}[\tilde{B}(\vec{k}, \vec{k}')]. \quad (1-6.7)$$

This equation states that the net energy transfer per unit time from frequencies in the vicinity of  $\vec{k}'$  to those near  $\vec{k}$  is proportional to the imaginary part of the bispectrum of the velocity field. Thus, as stated earlier, the bispectrum tells one how energy moves in frequency space.

### 1-7 Bispectrum = Decomposition of Cube of Time Series.

Defining the power spectrum and the bispectrum as in section 1.2, one gets the Parseval relation

$$\langle y^2(t) \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} P(\omega) d\omega, \quad (1-7.1)$$

which shows that the power spectrum represents the contribution to the total energy (or second moment) of a particular frequency range. One also finds the relation

$$\langle y^3(t) \rangle = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B(\omega_1, \omega_2) d\omega_1 d\omega_2, \quad (1-7.2)$$

which shows that the bispectrum represents the contribution to the (unnormalized) *skewness*, or third moment, of a particular *pair* of frequencies.

This result demonstrates how deeply the analogy between the second and third order quantities persists and modestly contributes to the reader's insight into the interpretation of the bispectrum. It is difficult to attach much physical

<sup>3</sup>For completeness, the "vector" bispectrum in the equation above is defined as

$$\tilde{B}(\vec{k}, \vec{k}') = \sum_{n=1}^3 \frac{1}{2\pi^6} \int \tilde{R}_n(\vec{x}, \vec{x}') \exp(-i[\vec{k} \cdot \vec{x} + \vec{k}' \cdot \vec{x}']) d\vec{x} d\vec{x}'$$

where the "vector" triple correlation in turn is

$$\tilde{R}_n(\vec{r}, \vec{s}) = \langle v_I(\vec{x}) v_n(\vec{x} + \vec{r}) v_n(\vec{x} + \vec{s}) \rangle.$$

or mathematical importance to this result (which is why it has been relegated to the end of this chapter).

### 1-8 The Bispectrum Can Be Interpreted in a Variety of Ways.

This introductory chapter has presented quite a number of ways of regarding the bispectrum. It presents, in fact, a rather broad survey of the bispectral literature. With luck, at least one or two of these perspectives may benefit the reader. Some of these examples (particularly the homogeneous turbulence example) may be overly technical: if so, ignore them! It is certainly true that understanding these specific examples is not necessary in order to understand the bispectrum.

The remainder of this report is more concrete. The following chapter presents some of the fundamental mathematics of the bispectrum, including a derivation of the usually mysterious "fundamental domain". From this point on, all mathematics will either be worked out in detail or, at least, stated carefully and precisely.

## CHAPTER 2

### Mathematical Properties of the Bispectrum

#### 2-1 The Principal Domain of a Bispectrum is a Consequence of Stationarity, "Reality", and Symmetries.

This section will involve discrete time series solely. The Fourier transform is in this case a continuous function of frequency which is  $2\pi$  periodic. Given a real, stationary third-order correlation function  $c_3(n_1, n_2, n_3)$ , the principal domain of the bispectrum is as shown in Figure 2-1. This fundamental domain consists of an isosceles triangular subset and an *odd* or *extra* triangle.

The shape of this principal domain is a consequence of stationarity, reality, discrete time, and symmetry properties. This shall be worked out in the next few pages.

Recall that  $c_3$  is defined as

$$c_3(n_1, n_2, n_3) = \langle x(t + n_1)x(t + n_2)x(t + n_3) \rangle. \quad (2-1.1)$$

First, the discrete time nature of  $c_3$  instructs one to represent the Fourier transform as

$$f_3(\lambda_1, \lambda_2, \lambda_3) = \frac{1}{(2\pi)^3} \sum_{n_1, n_2, n_3 = -\infty}^{\infty} c_3(n_1, n_2, n_3) \exp[-i(\lambda_1 n_1 + \lambda_2 n_2 + \lambda_3 n_3)] \quad (2-1.2)$$

where  $-\pi < \lambda_1, \lambda_2, \lambda_3 \leq \pi$  [2-1].

Second, one imposes stationarity on  $c_3$ ; i.e., one requires that

$$c_3(n_1, n_2, n_3) = c_3(n_1 - n_3, n_2 - n_3, 0) \equiv C_3(n_1 - n_3, n_2 - n_3). \quad (2-1.3)$$

Here the last equality serves as definition of the usual form for the stationary bicoherence function. For the sake of completeness, note that this function obeys

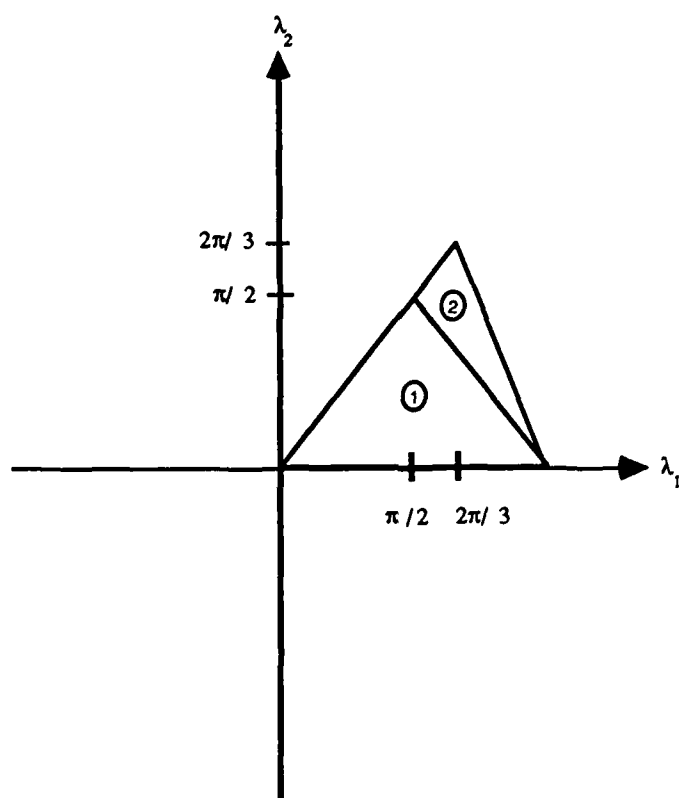


Figure 2-1 FUNDAMENTAL DOMAIN OF THE BISPECTRUM.

the symmetries

$$\begin{aligned} C_3(n_1, n_2) &= C_3(n_2, n_1) = \\ C_3(-n_1, n_2 - n_1) &= C_3(n_1 - n_2, -n_2). \end{aligned} \quad (2-1.4)$$

But, more relevantly, this stationarity implies that  $f_3$  is nonzero only when

$$(\lambda_1 + \lambda_2 + \lambda_3) \bmod 2\pi = 0. \quad (2-1.5)$$

One can now write the bispectrum as

$$B(\omega_1, \omega_2) = \langle X(\omega_1)X(\omega_2)X(\omega_3) \rangle, \quad (2-1.6)$$

where

$$\omega_3 = 2\pi n - \omega_1 - \omega_2, \quad (2-1.7)$$

and the value of  $n$  is restricted to 0 and 1 when  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  are confined to the range  $-\pi$  to  $\pi$ . (The value  $n = -1$  is also permissible but it gives redundant information and will be eliminated.)

At first sight, one might expect that, at least for some  $\omega_1$ ,  $\omega_2$ , there may be two possible values of  $\omega_3$  (one corresponding to  $n = 0$ , the other to  $n = 1$ ). In fact, for any specified  $\omega_1$  and  $\omega_2$  there is a unique  $n$  as well as a unique  $\omega_3$ .

It is clear that the value of the bispectrum is not changed if any pair of frequencies is interchanged. This observation provides the following symmetries

$$\begin{aligned} B(\omega_1, \omega_2) &= B(\omega_2, \omega_1), \\ B(\omega_1, \omega_2) &= B(\omega_3, \omega_1), \quad \text{and} \\ B(\omega_1, \omega_2) &= B(\omega_3, \omega_2), \end{aligned} \quad (2-1.8)$$

which are depicted in Figure 2-2.

For  $n = 0$ , these equations become

$$\begin{aligned} B(\omega_1, \omega_2) &= B(-\omega_1 - \omega_2, \omega_1) \quad \text{and} \\ B(\omega_1, \omega_2) &= B(-\omega_1 - \omega_2, \omega_2), \end{aligned} \quad (2-1.9)$$

as shown in Figure 2-3.

For  $n = 1$ , these relationships are

$$B(\omega_1, \omega_2) = B(2\pi - \omega_1 - \omega_2, \omega_1) \quad \text{and} \quad (2-1.10)$$

$$B(\omega_1, \omega_2) = B(2\pi - \omega_1 - \omega_2, \omega_2), \quad (2-1.11)$$

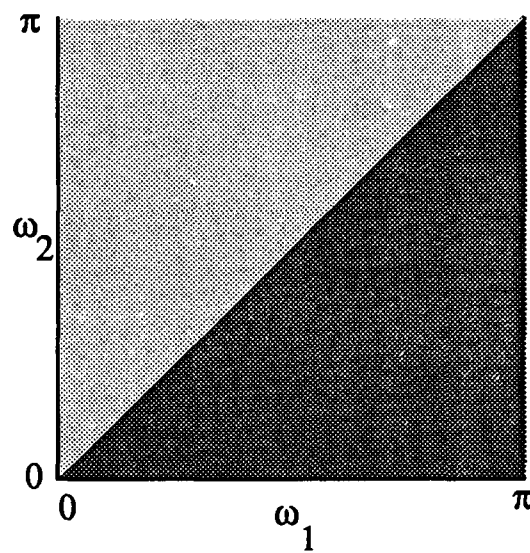


Figure 2-2  $B(\omega_1, \omega_2) = B(\omega_2, \omega_1)$ .

AS-88-210

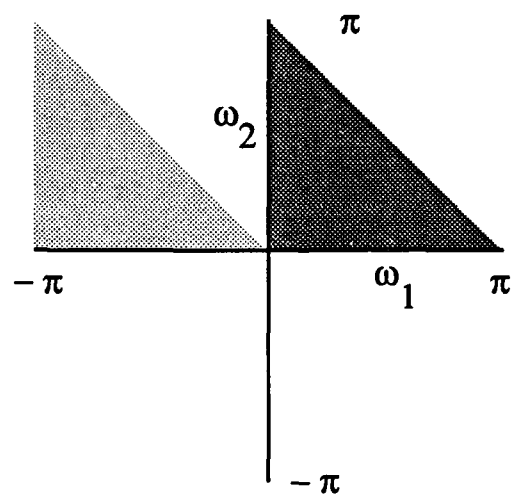


Figure 2-3  $B(\omega_1, \omega_2) = B(-\omega_1 - \omega_2, \omega_2)$ .

AS-88-211

as shown in Figures 2-4 and 2-5.

If the original time series is assumed to be real the bispectrum possesses the additional symmetry

$$B(\omega_1, \omega_2) = [B(-\omega_1, -\omega_2)]^* , \quad (2-1.12)$$

where the asterisk denotes complex conjugation. This is shown in Figure 2-6.

Using these symmetries, a principal domain of the bispectrum can be identified. From the conjugation symmetry, it follows that the regions of the  $(\omega_1, \omega_2)$  plane shown in Figure 2-6 are equivalent. Thus it is only necessary to consider the upper half of the plane (positive  $\omega_2$ ). From the second  $n = 0$  equation, the two triangular regions shown in Figure 2-3 are equivalent. Combining the second  $n = 0$  equation with the conjugation symmetry and the  $\omega_1 \leftrightarrow \omega_2$  symmetry, it can be seen that the two remaining triangular regions are also equivalent so that it is only necessary to consider positive values of both  $\omega_1$  and  $\omega_2$ . The  $\omega_1 \leftrightarrow \omega_2$  symmetry applied to the positive  $(\omega_1, \omega_2)$  quadrant results in the equivalence of the two octants shown in Figure 2-2. The first  $n = 1$  equation implies that the two triangles in Figure 2-4 are equivalent, while the second  $n = 1$  equation implies that the two triangles in Figure 2-5 are equivalent. No other symmetry relationships can be applied to further eliminate equivalent regions in the bispectrum. Thus the principal domain is the triangular region shown in Figure 2-1.

## 2-2 For a Properly Sampled Continuous Time Signal, the "Extra" Triangle is "Forbidden".

Data processing with digital computers necessitates dealing with discrete (*sampled*) representations of signals. Thus the bulk of this report emphasizes the discrete time case. However, time (so far as one knows) is fundamentally continuous. Moreover, the continuous nature of time makes itself felt even in discrete time signal processing. Specifically, if (1) the discrete time model arises from sampling an underlying continuous time process and (2) this continuous time process is bandlimited and (3) the sampling rate is sufficient to avoid aliasing, then the discrete bispectrum is nonzero only in the isosceles triangular subset of the principal domain.

This result shall be shown below.

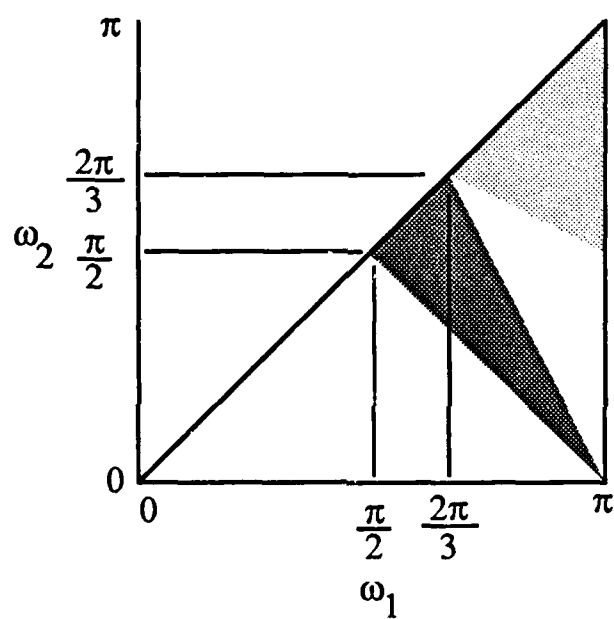


Figure 2-4  $B(\omega_1, \omega_2) = B(2\pi - \omega_1 - \omega_2, \omega_1)$ .



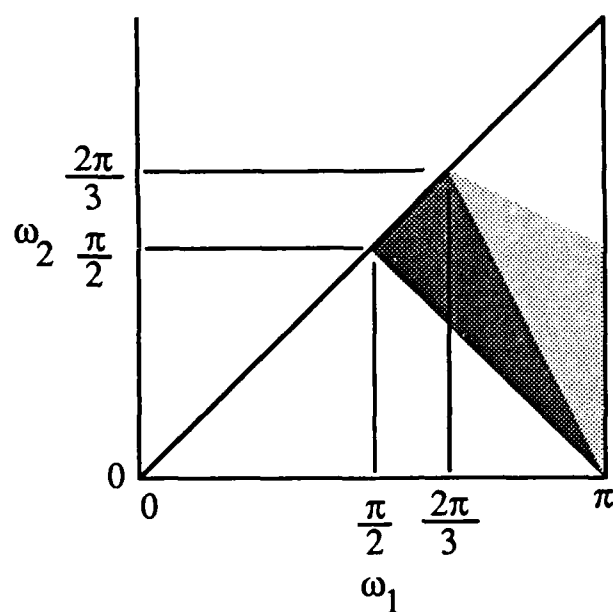


Figure 2-5  $B(\omega_1, \omega_2) = B(2\pi - \omega_1 - \omega_2, \omega_2)$ .

AS-88-213

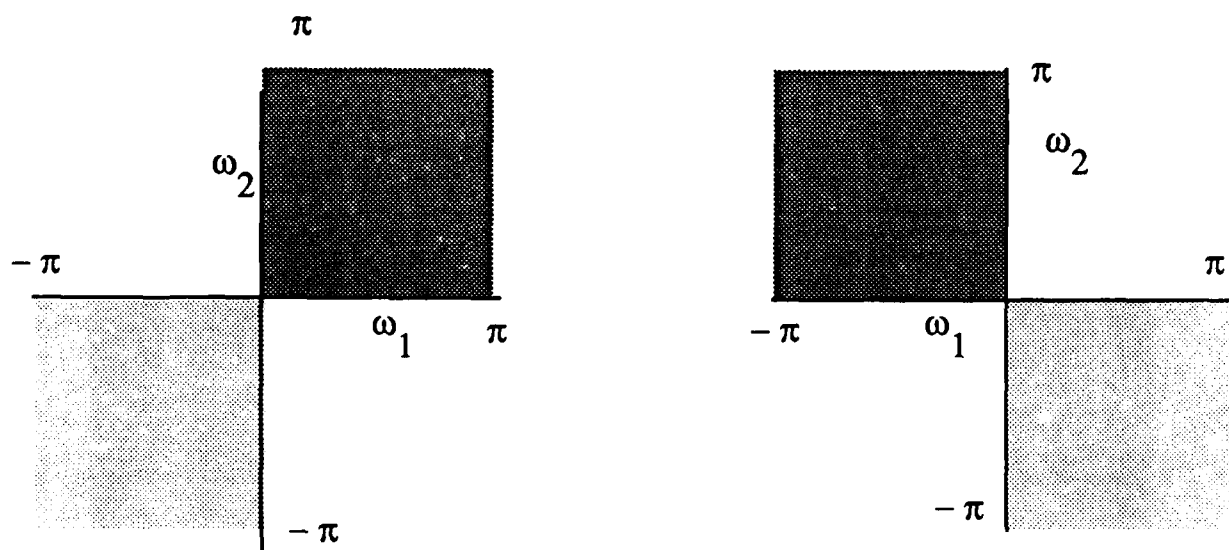


Figure 2-6  $B(\omega_1, \omega_2) = [B(-\omega_1, -\omega_2)]^*$ .

AS-88-214

A "bridge" between the discrete time bispectrum,  $\tilde{b}(\lambda_1, \lambda_2, \lambda_3)$ , and the continuous time bispectrum,  $b(\omega_1, \omega_2, \omega_3)$ , is given in Brillinger and Rosenblatt [2-2].

$$\tilde{b}(\lambda_1, \lambda_2, \lambda_3) = \sum_{\omega_1, \omega_2, \omega_3} b(\omega_1, \omega_2, \omega_3) \quad (2-2.1)$$

where  $\lambda_1 + \lambda_2 + \lambda_3 = 0 \pmod{2\pi}$ , and  $\omega_1 + \omega_2 + \omega_3 = 0$  with  $\omega_i = \lambda_i + 2\pi j_i$  and  $j_i$  is an integer (and  $i = 1, 2$ , or  $3$ ). (In this section, the tilde is used to denote *discrete time*.)

The bispectra are written with the third argument explicit for simplicity of exposition. Notice that the arguments of  $b$  must sum to zero whereas the arguments of  $\tilde{b}$  can be any integral multiple of  $2\pi$ . Also notice that this sum is consistent with the *replication* phenomenon familiar in sampled systems. That is, the continuous-time spectrum is replicated an infinite number of times over a three-dimensional lattice with side length  $2\pi$ .

Now assume that the underlying system is bandlimited. Further assume that the system has been sampled appropriately (above the Nyquist frequency). Then  $\pi$  corresponds to the Nyquist frequency or above and one may state that

$$b(\omega_1, \omega_2, \omega_3) = 0, \text{ if } |\omega_1|, |\omega_2|, \text{ or } |\omega_3| > \pi. \quad (2-2.2)$$

Now the argument is simple. Choose a point in the odd triangle. This implies that

$$\lambda_1 + \lambda_2 + \lambda_3 = 2\pi$$

and there is no reason not to take

$$-\pi \leq \lambda_i < \pi.$$

All  $\omega_1, \omega_2, \omega_3$  which correspond to this set of  $\lambda$  are included in the equation

$$0 = \omega_1 + \omega_2 + \omega_3 = \lambda_1 + \lambda_2 + \lambda_3 + 2\pi(j_1 + j_2 + j_3), \quad (2-2.3)$$

where the  $j_i$  are all integers (possibly zero). It is immediately clear that one has

$$j_1 + j_2 + j_3 = -1. \quad (2-2.4)$$

Thus at least one of the  $j_i$  cannot be zero. The simplest case is  $j_1 = -1$  and  $j_2$  and  $j_3 = 0$ . This leads to

$$\begin{aligned} \omega_1 &= \lambda_1 - 2\pi \\ \omega_2 &= \lambda_2 \\ \omega_3 &= \lambda_3, \end{aligned} \quad (2-2.5)$$

so  $\omega_1$  must lie outside of the Nyquist range. But this is a contradiction and shows that a properly sampled signal cannot lead to a nonzero bispectrum in the odd triangle.

Although this result is not surprising, the relevant literature has been ambivalent regarding its validity. Part of the reason for skepticism is the existence of data taken under conditions where aliasing is very unlikely but which shows unmistakable peaks in the *forbidden* triangle. The author's conclusion is that the signals involved must have been *non-stationary*, and that it would be desirable to pursue an understanding of the bispectrum in the non-stationary case. (The only mention of this extension in the literature that the author is aware of is to a problem involving speech recognition.)

### 2-3 The Most Basic Mathematical Properties of the Bispectrum are those which Determine the Fundamental Domain.

The elementary properties of the bispectrum, namely its symmetries and its relation to a stationary time series, are those properties which determine the fundamental domain. As shown here, however, this derivation requires working out the combined implications of all of these properties simultaneously. The treatment here is perhaps one of the most explicit available.

The next section shall extend the mathematical study of the bispectrum. The study will be done in the context of a concrete model which can be used to reproduce any possible bispectrum. This model should provide the reader with both insight and some useful manipulative techniques for studying the bispectrum.

## CHAPTER 3

### The Universal Bispectrum Model

#### 3-1 The Universal Bispectrum Model is the Simplest Model that can Reproduce any Possible Bispectrum.

This chapter thoroughly explores a particular (discrete) time series model which can reproduce the bispectrum of any possible time series. This model is not motivated by physical or intuitive considerations (unfortunately); however, it does provide a tool for learning useful mathematics, for testing software, and for illustrating and exploring the properties of the bispectrum.

It may be wise to start with the corresponding model for the power spectrum. Statisticians are very familiar with the time series model

$$x(t) = \epsilon(t) + \sum_{m=0}^{\infty} g(m)\epsilon(t-m). \quad (3-1.1)$$

This model represents an (infinite order) moving average (MA) process. Here  $\epsilon(t)$  is a random variable (taken to be normally distributed, for simplicity) and values of  $\epsilon$  at different times are statistically independent. In standard notation, the  $\epsilon(t) = N(0, \sigma^2)$  and are "independently, identically distributed" (i.i.d.). The  $\epsilon(t)$  can be regarded as random "inputs" (or *innovations*). The coefficients (or *weights*)  $g(m)$  can be chosen such that the time series  $x(t)$  has any desired power spectrum. So, this model may be considered to be a *universal* power spectrum model. (There are other models with this property.)

The *universal* bispectrum model is a simple generalization of the above. A convenient form is

$$x(t) = \epsilon(t) + \eta(t) + \sum_{\substack{m=0 \\ n=0}}^{\infty} g(m, n)\epsilon(t-m)\eta(t-m-n), \quad (3-1.2)$$

where now the model contains two independent random processes. That is, for all values of  $t$ ,  $\epsilon(t)$  and  $\eta(t) = N(0, \sigma^2)$  are i.i.d. It may suffice for now to think of

the universal bispectrum model as a sort of infinite MA model. The  $g(m, n)\epsilon(t)$  can be regarded as random MA weights and the  $\eta(t)$  can be regarded as the innovations. The model is completely specified by  $\sigma$  and the weights  $g$ . There are other models which can produce any possible bispectrum, but the above is the simplest to use.

In an initial reading, the remainder of this chapter can be omitted. Its primary purpose is to document the universal bispectrum model for future use. Necessarily it is very mathematical. However the mathematics have been spelled out in perhaps painful detail so as to be appropriate for an *introduction* to the bispectrum. The hope is that this chapter can serve to make the earlier discussions more concrete and thus enhance the reader's ability to deal with the bispectrum productively. The remaining sections of this chapter are organized as follows: the time domain properties of this model – the second-order (3.2) and third-order (3.3) cumulants are derived; these functions are then Fourier transformed to get the power spectrum (3.4) and the bispectrum (3.5); the inverse problem is examined (3.6); the model is reformulated and studied in the frequency domain (3.7,8); and finally, speculations on ways in which this model may arise in practice (3.9) are presented.

### 3-2 The Second-order Cumulant for the Universal Bispectrum Model is Rather Messy and Not Very Informative.

First on the agenda is to obtain the second-order cumulant function:

$$\begin{aligned}
 c_2(r) &\equiv \langle x(t) x(t+r) \rangle \\
 &= \langle [\epsilon(t) + \sum_{m,n} g(m,n)\epsilon(t-m)\eta(t-m-n)] \\
 &\quad [\epsilon(t+r) + \eta(t+r) \\
 &\quad + \sum_{m',n'} g(m',n')\epsilon(t+r-m')\eta(t+r-m'-n')] \rangle \quad (3-2.1) \\
 &= \langle \epsilon(t)\epsilon(t+r) + \eta(t)\eta(t+r) \\
 &\quad + \sum_{m,m',n,n'} g(m,n)g(m',n') \\
 &\quad \epsilon(t-m)\epsilon(t+r-m')\eta(t-m-n)\eta(t+r-m'-n') \rangle.
 \end{aligned}$$

So far we have just substituted for  $x$  and used the fact that  $\epsilon$  and  $\eta$  are independent variables. Next we shall use the fact that  $\epsilon(t)$  is independent of all  $\epsilon(t')$  for  $t' \neq t$  (and similarly for  $\eta$ ). The expectation of the  $\epsilon(t)\epsilon(t+r)$  term becomes  $\sigma^2\delta_r$  (and

so does the corresponding term for  $\eta$ ). Here  $\delta_r$  is the *Kronecker* delta function which is zero except when its argument ( $r$ ) has the value zero, in which case it is one. The expectation of the complicated summation also simplifies in so far as only those terms for which  $m' = m + r$  and  $n' = n$  contribute. So the final expression becomes

$$c_2(r) = 2\sigma^2\delta_r + \sigma^4 \sum_{m,n=0}^{\infty} g(m,n)g(m+r,n). \quad (3-2.2)$$

This expression is included more for completeness than for its usefulness. Note that this formula does have the  $r \leftrightarrow -r$  symmetry of  $c_2(r)$ , but that this symmetry is not blatantly on display here.

### 3-3 For the Universal Bispectrum Model, the Third-Order Cumulant Equals $\sigma^4$ times the Weight.

The third-order cumulant function is evaluated in a similar fashion (but it turns out nicer):

$$\begin{aligned} c_3(r,s) &\equiv \langle x(t)x(t+r)x(t+s) \rangle = \\ &\quad \underbrace{1}_{\epsilon(t)} + \underbrace{2}_{\eta(t)} + \underbrace{3}_{\sum_{m,n} g(m,n)\epsilon(t-m)\eta(t-m-n)} \quad (3-3.1) \\ &\quad \underbrace{A}_{\epsilon(t+r)} + \underbrace{B}_{\eta(t+r)} + \underbrace{C}_{\sum_{m',n'} g(m',n')\epsilon(t+r-m')\eta(t+r-m'-n')} \\ &\quad \underbrace{i}_{\epsilon(t+s)} + \underbrace{ii}_{\eta(t+s)} + \underbrace{iii}_{\sum_{m'',n''} g(m'',n'')\epsilon(t+s-m'')\eta(t+s-m''-n'')} \end{aligned}$$

The above product contains 27 factors. However the single rule that only terms with an even number of factors of  $\epsilon$ 's and  $\eta$ 's will have a non-vanishing expectation eliminates all but six terms. These six terms all involve a (double)sum of a product of a  $g$ , two  $\epsilon$ 's, and two  $\eta$ 's. In order for these terms to contribute, the arguments of the two  $\epsilon$ 's must be at the same time and similarly for the  $\eta$ 's. This reduces each of the six double summations to a single term. These six terms are shown below using the notation *source*  $\rightarrow$  *target*. The *source* specifies the particular combination of three factors which leads to a non-vanishing expectation and the *target* denotes the weight  $g$  involved in the surviving combination.

$$\begin{aligned}
c_3(r, s) &\equiv \langle x(t)x(t+r)x(t+s) \rangle \\
&= \langle \quad 1Biii \longrightarrow g(s, -r) \\
&\quad + 1Cii \longrightarrow g(r, -s) \\
&\quad + 2Aiii \longrightarrow g(s-r, r) \\
&\quad + 2Ci \longrightarrow g(r-s, s) \\
&\quad + 3Aii \longrightarrow g(-r, r-s) \\
&\quad + 3Bi \longrightarrow g(-s, s-r) \rangle. \tag{3-3.2}
\end{aligned}$$

The two  $\epsilon$ 's and  $\eta$ 's in these terms contribute a factor of  $\sigma^4$  to the expectation above, so the final result is

$$\begin{aligned}
c_3(r, s) = \sigma^4 \quad \{ \quad &g(s, -r) \\
&+ g(r, -s) \\
&+ g(s-r, r) \\
&+ g(r-s, s) \\
&+ g(-r, r-s) \quad \}. \tag{3-3.3}
\end{aligned}$$

This equation is already much nicer than equation (3-2.2), but it is even better than it appears. For a specific choice of  $r$  and  $s$ , in general only one term will be non-zero in the above sum. For example, if  $r$  and  $s$  are both positive, and if  $r > s$ , then only the  $g(r-s, s)$  term is non-zero. *Causality* (the fact that  $g(r, s) = 0$  if  $r < 0$  or  $s < 0$ ) reduces the above sum to one term. Moreover, the remaining terms are just those that are needed for  $c_3(r, s)$  to obey the symmetry requirements of a third-order correlation function. So, in essence, the third-order correlation function  $c_3(r, s)$  is equal to  $g(r-s, s)$ , except for a factor of  $\sigma^4$ , and unless  $g(r-s, s) = 0$ , in which case one must take the appropriate non-zero image.

However, things are slightly messier than implied above. If either  $r$  or  $s$  is equal to zero, or if  $r$  and  $s$  are equal to one another, then several terms in the above sum may simultaneously be non-zero, so that  $c_3(r, s)$  and  $g(r-s, s)$  differ by a small factor in addition.

The exact relation, which is equivalent to equation (3-3.3), is

$$c_3(r, s) = \sigma^4 \left\{ \begin{array}{lll} 6g(0, 0) & r = 0, s = 0 & \boxed{0} \\ 2g(r, 0) & r > 0, s = 0 & \boxed{1} \\ g(r - s, s) & r > s, s > 0 & \boxed{2} \\ 2g(0, r) & r = s, s > 0 & \boxed{3} \\ g(s - r, r) & r > 0, s > r & \boxed{4} \\ 2g(0, s) & r = 0, s > 0 & \boxed{5} \\ g(s, -r) & r < 0, s > 0 & \boxed{6} \\ 2g(0, -r) & r < 0, s = 0 & \boxed{7} \\ g(-s, s - r) & r < s, s < 0 & \boxed{8} \\ 2g(-r, 0) & r = s, s < 0 & \boxed{9} \\ g(-r, r - s) & r < 0, s < r & \boxed{10} \\ 2g(0, -s) & r = 0, s < 0 & \boxed{11} \\ g(r, -s) & r > 0, s < 0 & \boxed{12} \end{array} \right. \quad (3-3.4)$$

The boxed numbers ranging from 0 to 12 refer to Figure 3-1.

This figure presents the  $r, s$  plane and shows how each term above contributes. Each region assigned an odd number corresponds to the line which is the boundary between the two adjacent wedges. Each region assigned an even number represents one of the six primary regions above (except for the region assigned the number zero, which represents the origin only). Thus, for example, in region  $\boxed{6}$ ,  $c_3(r, s)$  is equal to  $g(s, -r)$ .

### 3-4 The Power Spectrum of the Universal Bispectrum Model is Somewhat Informative.

The power spectrum is defined as the Fourier transform of the second-order cumulant function, so that

$$\begin{aligned} f_2(\lambda) &= \sum_{r=-\infty}^{\infty} c_2(r) e^{-i\lambda r} \\ &= \sum_{r=-\infty}^{\infty} [2\sigma^2\delta_r + \sigma^4 \sum_{m,n} g(m, n)g(m + r, n)] e^{-i\lambda r} \end{aligned} \quad (3-4.1)$$



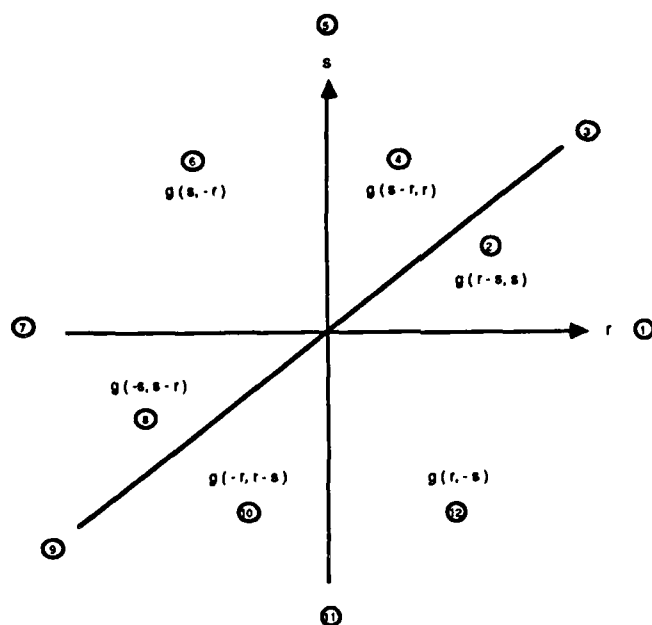


Figure 3-1 THE  $c_3$  PLANE IN TERMS OF THE WEIGHTS  $g$ .

$$\begin{aligned}
 &= 2\sigma^2 + \sigma^4 \sum_{m,n} g(m,n) \sum_{r=-\infty}^{\infty} g(m+r,n) e^{-i\lambda r} \\
 &= 2\sigma^2 + \sigma^4 \sum_{m,n} g(m,n) \sum_{s=-\infty}^{\infty} g(s,n) e^{-i\lambda(s-m)}. \quad (3-4.2)
 \end{aligned}$$

Therefore, if one defines  $\hat{g}_n(\lambda) \equiv \sum_{k=-\infty}^{\infty} g(k,n) e^{-i\lambda k}$ , then

$$\boxed{f_2(\lambda) = 2\sigma^2 + \sigma^4 \sum_{n=0}^{\infty} |\hat{g}_n(\lambda)|^2.} \quad (3-4.3)$$

This result is somewhat helpful in limiting circumstances. In particular, if the weights are small in comparison with  $\sigma$ , the power spectrum will be approximately flat (independent of frequency) with magnitude  $2\sigma^2$ . In the opposite limit, the power spectrum will vary as the magnitude of a typical  $g$  squared.

This result will also be useful later in determining how to trade off the parameters ( $\sigma$  and the  $g$ 's) to yield the model with smallest variance for a specified bispectrum.

### 3-5 The Bispectrum of the Universal Bispectrum Model is Given by the Frequency Coupling Coefficients.

Remember that the bispectrum is the double Fourier transform of the third-order correlation function:

$$b(\lambda_1, \lambda_2) = \sum_{r,s=-\infty}^{\infty} c_3(r, s) e^{-i(\lambda_1 r + \lambda_2 s)}. \quad (3-5.1)$$

The following definition is needed in order to re-express  $b(\lambda_1, \lambda_2)$  in terms of the weights:

$$\hat{g}(\lambda_1, \lambda_2) = \sum_{r,s=-\infty}^{\infty} g(r, s) \exp[-i(r\lambda_1 + s\lambda_2)]. \quad (3-5.2)$$

The  $\hat{g}$  will be referred to as *frequency coupling constants* for reasons to be made clear in section 3.6. Substitute for  $c_3(r, s)$  using equation (3-3.3). The result is a sum of six terms. The detailed calculation for the term involving  $g(r - s, s)$  follows. The contribution of this term to the entire bispectrum will be denoted  $b_{\boxed{2}}$  because only  $g(r - s, s)$  is non-zero in region  $\boxed{2}$ , thus

$$\begin{aligned} b_{\boxed{2}}(\lambda_1, \lambda_2) &= \sigma^4 \sum_{r,s=-\infty}^{\infty} g(r - s, s) \exp[-i(r\lambda_1 + s\lambda_2)] \\ &= \sigma^4 \sum_{r',s=-\infty}^{\infty} g(r', s) \exp[-i(r'\lambda_1 + s(\lambda_1 + \lambda_2))] \\ &= \sigma^4 \hat{g}(\lambda_1, \lambda_1 + \lambda_2). \end{aligned} \quad (3-5.3)$$

Similarly,

$$\begin{aligned} b_{\boxed{4}}(\lambda_1, \lambda_2) &= \sigma^4 \hat{g}(\lambda_2, \lambda_1 + \lambda_2) \\ b_{\boxed{6}}(\lambda_1, \lambda_2) &= \sigma^4 \hat{g}(\lambda_2, -\lambda_1) \\ b_{\boxed{8}}(\lambda_1, \lambda_2) &= \sigma^4 \hat{g}(-\lambda_1 - \lambda_2, -\lambda_1) \\ b_{\boxed{10}}(\lambda_1, \lambda_2) &= \sigma^4 \hat{g}(-\lambda_1 - \lambda_2, -\lambda_2) \\ b_{\boxed{12}}(\lambda_1, \lambda_2) &= \sigma^4 \hat{g}(\lambda_1, -\lambda_2). \end{aligned} \quad (3-5.4)$$

The final result is just the sum of these six terms:

$$b(\lambda_1, \lambda_2) = \sigma^4 [ \hat{g}(\lambda_1, \lambda_1 + \lambda_2) + \hat{g}(\lambda_2, \lambda_1 + \lambda_2) + \hat{g}(\lambda_2, -\lambda_1) + \hat{g}(-\lambda_1 - \lambda_2, -\lambda_1) + \hat{g}(-\lambda_1 - \lambda_2, -\lambda_2) + \hat{g}(\lambda_1, -\lambda_2) ]$$

$$\begin{aligned}
& + \hat{g}(\lambda_2, -\lambda_1) + \hat{g}(\lambda_1, -\lambda_2) \\
& + \hat{g}(-\lambda_1 - \lambda_2, -\lambda_1) + \hat{g}(-\lambda_1 - \lambda_2, -\lambda_2)]. \quad (3-5.5)
\end{aligned}$$

(The reader may have some doubt as to whether this result takes care of the "boundary" regions  $\boxed{0}$ ,  $\boxed{1}$ ,  $\boxed{3}$  etc., properly. Rest assured that it does. Notice that  $b_{\boxed{2}}$  includes one-sixth of the contribution from  $\boxed{0}$ , and one-half of the contribution from  $\boxed{1}$  and  $\boxed{3}$ , so that the sum of all six terms does in fact correctly include the regions which do not explicitly appear above.)

As was the case for the corresponding time domain result, this formula is much simpler than that for the second-order quantity. Further, it is simpler than it appears. Originally, the author used the above equation (3-5.5) to base a presentation of the "inverse" problem of constructing the time domain weights due to a particular bispectrum. The author currently favors the approach shown in the next section so that equation (3-5.5) is no longer so important to the development. Yet it is still of interest to solve the above equation for  $\hat{g}$  given  $b$  and the next section shall show how this can be done.

### 3-6 Inversion of the Model is Easy.

This section will show how to go backwards. That is, given a bispectrum, the corresponding model coefficients will be determined. For simplicity, it shall be assumed that the desired bispectrum lies entirely in the interior of the isosceles triangular subset of the principal domain. This assumption makes exposition somewhat easier, but it is by no means difficult to redo this section for the general case.

Given a bispectrum  $b$  as assumed above, define  $b_1(\omega_1, \omega_2)$  to be a  $2\pi$  periodic (in  $\omega_1, \omega_2$ ), conjugation symmetric function which agrees with  $b$  in the fundamental domain but is zero elsewhere. One can call  $b_1$  the signature of the bispectrum. The signature is  $b$  stripped of all the symmetries of  $b$  except for the periodicity and the conjugation symmetry. In terms of  $b_1$ ,  $b$  can be written

$$\begin{aligned}
b(\omega_1, \omega_2) = & b_1(\omega_1, \omega_2) \\
& + b_1(\omega_2, \omega_1) \\
& + b_1(\omega_1, -\omega_1 - \omega_2) \\
& + b_1(\omega_2, -\omega_1 - \omega_2) \\
& + b_1(-\omega_1 - \omega_2, \omega_1) \\
& + b_1(-\omega_1 - \omega_2, \omega_2)
\end{aligned}$$

+ c.c.

This expression shows that  $b$  is a symmetrized version of its signature and has all the correct bispectral symmetries. Thus  $b_1$  is, as claimed,  $b$  "de-symmetrized". The author derived this equation by starting with the  $b_1(\omega_1, \omega_2)$  piece and applying the bispectral symmetries to get the complete bispectrum. The assumption that  $b$  is non-zero in only the interior of the fundamental domain has already been used, for otherwise the above sum would overcount contributions on the boundary. This overcounting problem is, in fact, the reason for the simplifying assumption.

Assume next that the inverse double Fourier transform of  $b_1$  together with the conjugation-symmetric part of  $b_1$  is known. This double Fourier transform is real and it is defined by

$$h(r, s) = \frac{1}{4\pi^2} \int \int b_1(\omega_1, \omega_2) \exp[i(r\omega_1 + s\omega_2)] d\omega_1 d\omega_2 + c.c. \quad (3-6.1)$$

where the integration limits are  $-\pi$  to  $\pi$ .

It is a routine matter to verify the following table of transforms:

$$\begin{aligned} b_1(\omega_1, \omega_2) &\Rightarrow h(r, s) \\ b_1(\omega_2, \omega_1) &\Rightarrow h(s, r) \\ b_1(\omega_1, -\omega_1 - \omega_2) &\Rightarrow h(r - s, -s) \\ b_1(\omega_2, -\omega_1 - \omega_2) &\Rightarrow h(s - r, -r) \\ b_1(-\omega_1 - \omega_2, \omega_1) &\Rightarrow h(-s, r - s) \\ b_1(-\omega_1 - \omega_2, \omega_2) &\Rightarrow h(-r, s - r), \end{aligned} \quad (3-6.2)$$

where  $x \Rightarrow y$  symbolizes the fact that  $y$  is the inverse double Fourier transform of  $x$ .

Combining equations (3-6.1), (3-6.3), and the fact that the bispectrum is the double Fourier transform of the triple correlation gives

$$c_3(r, s) = h(r, s) + h(s, r) + h(r - s, -s) + h(s - r, -r) + h(-s, r - s) + h(-r, s - r). \quad (3-6.3)$$

This is a key result in the inversion of the bispectrum: it expresses the triple correlation in terms of the transform of the bispectral signature. This expression displays explicitly all the appropriate symmetries of a triple correlation function so that  $h$  like the  $g$  earlier is a de-symmetrized version of the triple correlation. All of this so far in this section is entirely independent of the universal

bispectrum model. But now, using the universal bispectrum model, one is essentially done. For given the triple correlation, there is a unique corresponding  $g$ . This  $g$  can be obtained by inverting (3-3.4) to get

$$g(r, s) = \frac{1}{\sigma^4} \begin{cases} \frac{1}{6} c_3(0, 0) & r = s = 0 \\ \frac{1}{2} c_3(r, 0) & r > 0, s = 0 \\ \frac{1}{2} c_3(0, s) & r = 0, s > 0 \\ c_3(r + s, s) & r, s > 0 \end{cases} \quad (3-6.4)$$

Actually the universal bispectrum model permits an even more direct solution to the inverse problem. It is not necessary to introduce the signature at all. One just takes the above equation and rewrites it as

$$\begin{aligned} g(r, s) &= \frac{1}{\sigma^4} \{c_3(r + s, s)[u(r)u(s) - \frac{1}{2} u(r)\delta(s) - \frac{1}{2} \delta(r)u(s) + \frac{1}{6} \delta(r)\delta(s)]\} \\ &= \frac{1}{\sigma^4} c'_3(r, s)m(r, s). \end{aligned} \quad (3-6.5)$$

In the first line above,  $u(r)$  is the step function ( $u(r) = 1$  for  $r \geq 0$  and  $u(r) = 0$  otherwise) and  $\delta(r)$  is the Kronecker delta function. In the second line  $c'_3(r, s) \equiv c_3(r + s, s)$  and  $m(r, s)$  is the sum of products of  $\delta$  and  $u$  bracketed above.

If one double Fourier transforms both sides of the above equation, then the left side will give the  $\hat{g}$ 's and the right side will be the convolution of the given bispectrum with the transform of  $m$ . Explicitly, one can write

$$\begin{aligned} \hat{g}(\lambda_1, \lambda_2) &= \frac{1}{\sigma^4} [\hat{c}'_3 \star \hat{m}(\lambda_1, \lambda_2)] \\ &\text{with } \hat{c}'_3(\lambda_1, \lambda_2) = \hat{b}(\lambda_1, \lambda_2 - \lambda_1). \end{aligned} \quad (3-6.6)$$

The transforms of the step function and the delta function are simple to get; they are

$$\begin{aligned} \hat{u}(\lambda) &= \frac{1}{1 - \exp(-i\lambda)} \\ &= \frac{1}{2} + \sum_{m=-\infty}^{\infty} \delta(\lambda - 2\pi m) + \\ &\quad \frac{i}{2} \cot(\lambda/2) \end{aligned} \quad (3-6.7)$$

and

$$\hat{\delta}(\lambda) = 1, \quad (3-6.8)$$

and  $\hat{m}$  may readily be obtained by appropriately combining the above results.

Writing the above convolution somewhat more explicitly gives

$$\hat{g}(\lambda_1, \lambda_2) = \frac{1}{\sigma^4} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \hat{b}(\lambda_1 - \omega_1, \lambda_2 - \lambda_1 - \omega_2) \hat{m}(\omega_1, \omega_2) d\omega_1 d\omega_2. \quad (3-6.9)$$

The above formula expresses the  $\hat{g}$ 's directly in terms of the given bispectrum. It is thus the inverse of the relation obtained at the end of the previous section. The expression above is not suited to numerical computations, however, due to the singularity of the cotangent functions at  $\omega = 0$ . Presumably it may be simplified using principal value techniques by analogy to the continuous time situation.

One could now go on to recover the  $g$ 's from the  $\hat{g}$ 's by an inverse double transform. This provides a logically straightforward direct route from the desired bispectrum to the model coefficients. Numerically, however, it is simpler merely to compute  $c_3$  via the double transform and then apply equation (3-6.5). The indirect route presented first is just an attempt to explicitly account for the symmetries of the bispectrum after doing the double transform.

### 3-7 The Universal Bispectrum Model in the Frequency Domain is as Simple as it Can Possibly Be.

The preceding sections have formulated the universal bispectrum model in the time domain (3.1), computed its time domain properties, first (3.2) and second (3.3) order cumulants, and then transformed these quantities into the frequency domain to find its power spectrum (3.4) and its bispectrum (3.5). An alternate technique is to formulate the model in the frequency domain and then to compute its power spectrum and its bispectrum directly. This approach will be taken in the present section. Here the "frequency coupling coefficients" take a more central role.

Start with the time domain version of the model, equation (3-1.2):

$$x(t) = \epsilon(t) + \eta(t) + \sum_{m,n=0}^{\infty} g(m,n) \epsilon(t-m) \eta(t-m-n). \quad (3-7.1)$$

Express  $x$ ,  $\epsilon$ ,  $\eta$  in terms of their Fourier components:

$$x(t) = \frac{1}{2\pi} \int_{2\pi} \hat{x}(\omega) e^{i\omega t} d\omega \quad (3-7.2)$$

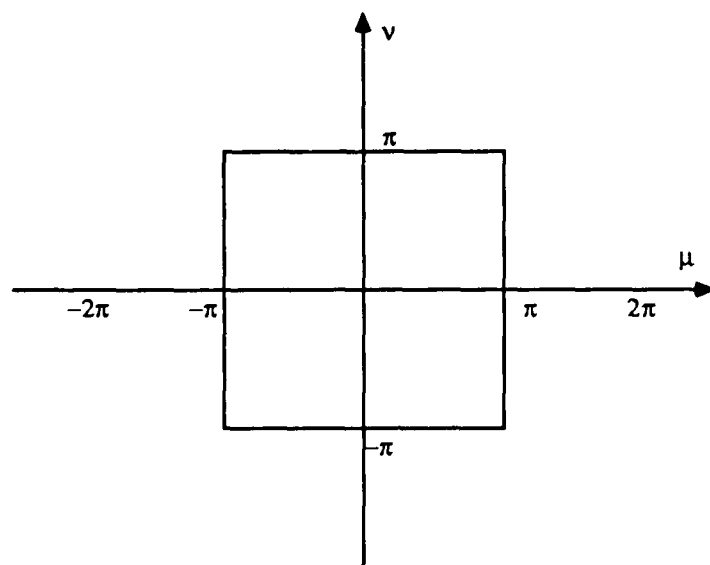


Figure 3-2 INTEGRATION LIMITS BEFORE CHANGE OF VARIABLES.

(with similar expressions for  $\hat{\epsilon}$  and  $\hat{\eta}$ ).

Therefore,

$$\begin{aligned} \int \hat{x}(\omega) e^{i\omega t} d\omega &= \int \hat{\epsilon}(\mu) e^{i\mu t} d\mu + \int \hat{\eta}(\omega) e^{i\omega t} d\omega \\ &+ \sum_{m,n=0}^{\infty} g(m,n) \frac{1}{2\pi} \int \hat{\epsilon}(\mu) e^{i\mu(t-m)} d\mu \int \hat{\eta}(\nu) e^{i\nu(t-m-n)} d\nu. \end{aligned} \quad (3-7.3)$$

Continuing, one gets

$$\begin{aligned} \int [\hat{x}(\omega) - \hat{\epsilon}(\omega) - \hat{\eta}(\omega)] e^{i\omega t} d\omega &= \frac{1}{2\pi} \sum_{m,n=0}^{\infty} g(m,n) \\ &\int \int \hat{\epsilon}(\mu) \hat{\eta}(\nu) e^{i(\mu+\nu)(t-m)} e^{-i\nu n} d\mu d\nu. \end{aligned} \quad (3-7.4)$$

Now is a good time to change variables from  $\mu, \nu$  to  $\omega = \mu + \nu, \nu$ . The Jacobian of this transformation is one (1), conveniently!. What about integration limits? The original region of integration in the  $\mu, \nu$  plane can be chosen as the square shown in Figure 3-2.

The new region is then the rhombus (Figure 3-3).

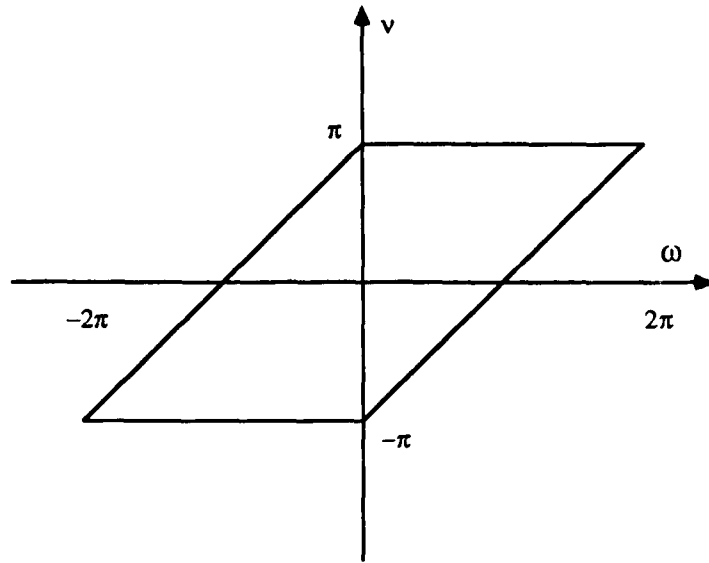


Figure 3-3 INTEGRATION LIMITS AFTER CHANGE OF VARIABLES.

Continuing,

$$\begin{aligned} \int \int [\hat{x}(\omega) - \hat{\epsilon}(\omega) - \hat{\eta}(\omega)] e^{i\omega t} d\omega = \\ \frac{1}{2\pi} \int \int \hat{\epsilon}(\omega - \nu) \hat{\eta}(\nu) \sum_{m,n=0}^{\infty} g(m,n) e^{-i(\omega m + \nu n)} e^{i\omega t} d\nu d\omega. \end{aligned} \quad (3-7.5)$$

So one is led to re-introduce the frequency coupling coefficients

$$\hat{g}(\omega, \nu) = \sum_{m,n=0}^{\infty} g(m,n) e^{-i(\omega m + \nu n)}. \quad (3-7.6)$$

Therefore,

$$\int [\hat{x}(\omega) - \hat{\epsilon}(\omega) - \hat{\eta}(\omega)] e^{i\omega t} d\omega = \frac{1}{2\pi} \int \int \hat{g}(\omega, \nu) \hat{\epsilon}(\omega - \nu) \hat{\eta}(\nu) e^{i\omega t} d\nu d\omega. \quad (3-7.7)$$

Further,

$$\int [\hat{x}(\omega) - \hat{\epsilon}(\omega) - \hat{\eta}(\omega)] e^{i\omega t} d\omega = \frac{1}{2\pi} \int \int \hat{g}(\omega, \nu) \hat{\epsilon}(\omega - \nu) \hat{\eta}(\nu) e^{i\omega t} d\nu d\omega. \quad (3-7.8)$$

Currently the integration domain is the rhombus. This can be replaced with the more convenient (original) square in the  $\omega, \nu$  plane by the following argument. Consider the triangular segment of the domain shown in Figure 3-4.



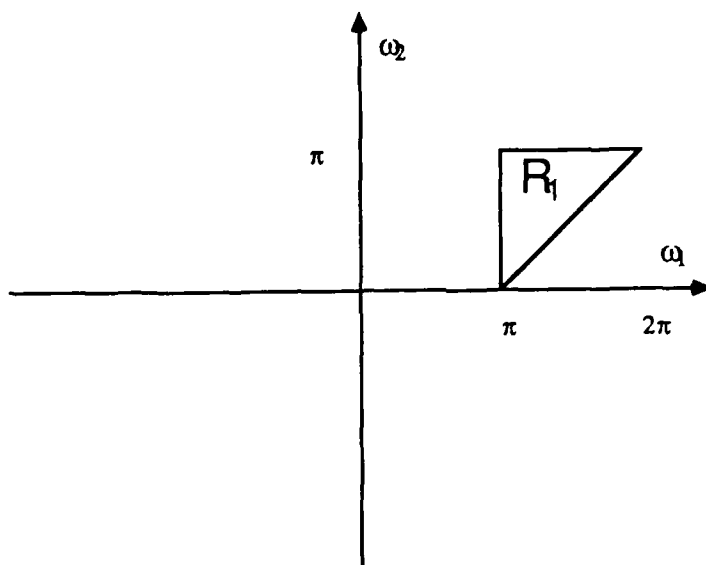


Figure 3-4 MIGRATING TRIANGULAR PIECE (BEFORE MIGRATION).

This piece contributes

$$\int \int_{R_1} \hat{g}(\omega, \nu) \hat{e}(\omega - \nu) \hat{\eta}(\nu) e^{i\omega t} d\nu d\omega. \quad (3-7.9)$$

Now change variables:  $\nu \rightarrow \nu$ ,  $\omega \rightarrow \omega' = \omega - 2\pi$ . One gets

$$\int \int_{R'_1} \hat{g}(\omega' + 2\pi, \nu) \hat{e}(\omega' + 2\pi - \nu) \hat{\eta}(\nu) e^{i(\omega' + 2\pi)t} d\nu d\omega'. \quad (3-7.10)$$

Using

$$\begin{aligned} \hat{g}(\omega, \nu) &= \hat{g}(\omega \pm 2\pi, \nu) = \hat{g}(\omega, \nu \pm 2\pi) \\ \hat{e}(\omega) &= \hat{e}(\omega \pm 2\pi), \end{aligned} \quad (3-7.11)$$

this becomes

$$\int \int_{R'_1} \hat{g}(\omega, \nu) \hat{e}(\omega - \nu) \hat{\eta}(\nu) e^{i\omega t} d\nu d\omega. \quad (3-7.12)$$

The triangle  $R_1$  has been moved over to become the triangle  $R'_1$  shown in Figure 3-5. In this fashion the region of integration can be restored to the original square.

The frequency domain equation becomes

$$\int [\hat{x}(\omega) - \hat{e}(\omega) - \hat{\eta}(\omega)] e^{i\omega t} d\omega = \int \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega, \nu) \hat{e}(\omega - \nu) \hat{\eta}(\nu) e^{i\omega t} d\nu \right\} e^{i\omega t} d\omega, \quad (3-7.13)$$

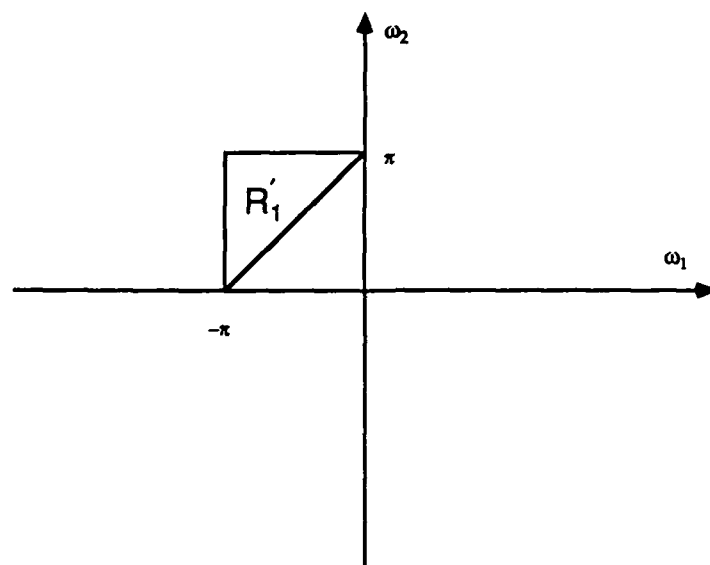


Figure 3-5 MIGRATING TRIANGULAR PIECE (AFTER MIGRATION).

so that the model is

$$\hat{x}(\omega) = \hat{e}(\omega) + \hat{\eta}(\omega) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega, \nu) \hat{e}(\omega - \nu) \hat{\eta}(\nu) d\nu. \quad (3-7.14)$$

Equivalently, in Cramér notation (described more fully in the next section),

$$d\hat{x}(\omega) = d\hat{e}(\omega) + d\hat{\eta}(\omega) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega, \nu) d\hat{e}(\omega - \nu) d\hat{\eta}(\nu). \quad (3-7.15)$$

### 3-8 The Bispectrum Can Be Computed Directly in the Frequency Domain.

It is now a simple calculation to get the bispectrum directly. (The calculation of the power spectrum is left as "an exercise for the reader".) Cramér notation makes the calculation much less messy. MacDonald's paper [1.4-2] presents a good discussion of the physical interpretation of the Cramér representation.

Let  $x(t)$  be a real, stationary, discrete time stochastic time series. Then

its Cramér representation is

$$x(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega t} d\hat{x}(\omega). \quad (3-8.1)$$

Useful properties are

$$\begin{aligned} d\hat{x}(-\omega) &= \overline{d\hat{x}(\omega)} \\ \text{and } \langle d\hat{x}(\omega_1) d\hat{x}(\omega_2) \rangle &= \delta'(\omega_1 + \omega_2) f_2(\omega_1) d\omega_1 d\omega_2 \end{aligned} \quad (3-8.2)$$

where  $c_2(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega\tau} f_2(\omega) d\omega$ .

For the independent Gaussian inputs (*innovations*), one finds

$$\begin{aligned} \langle d\hat{\epsilon}(\omega_1) d\hat{\epsilon}(\omega_2) \rangle &= \sigma^2 \delta'(\omega_1 + \omega_2) d\omega_1 d\omega_2, \\ \langle d\hat{\eta}(\omega_1) d\hat{\eta}(\omega_2) \rangle &= \sigma^2 \delta'(\omega_1 + \omega_2) d\omega_1 d\omega_2, \\ \langle d\hat{\epsilon}(\omega_1) d\hat{\eta}(\omega_2) \rangle &= 0. \end{aligned} \quad (3-8.3)$$

Here we use the notation  $\delta'(x) \equiv \sum_{k=-\infty}^{\infty} \delta(x + 2\pi k)$ . This is the  $2\pi$  periodic extension of the Dirac delta function that Brillinger calls the "Dirac comb" and denotes  $\eta(x)$ .

MacDonald presents the following formula for the bispectrum using the Cramér representation:

$$B(f_1, f_2) df_1 df_2 \delta(f_1, f_2, f_3) = E\{d\hat{x}(f_1), d\hat{x}(f_2), d\hat{x}(f_3)\}. \quad (3-8.4)$$

This formula must be put in terms of radian frequency  $\omega$  as opposed to angular frequency  $f$ . This means one must distinguish between the function  $d\hat{x}_f(f)$  and  $d\hat{x}_\omega(\omega)$ . The previous  $d\hat{x}$ 's are actually  $d\hat{x}_\omega$ 's. The Cramér representation for  $x$  in terms of angular frequency is

$$x(t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi ft} d\hat{x}_f(f). \quad (3-8.5)$$

and this implies that ( $\omega = 2\pi f$ )

$$d\hat{x}_f(f) = d\hat{x}_\omega(\omega). \quad (3-8.6)$$

Further, for all positive  $a$

$$\delta(ax) = \frac{1}{a} \delta(x). \quad (3-8.7)$$

Using these results, MacDonald's equation becomes

$$B(\omega_1, \omega_2) d\omega_1 d\omega_2 \delta(\omega_1, \omega_2, \omega_3) = 2\pi E\{d\hat{x}(\omega_1), d\hat{x}(\omega_2), d\hat{x}(\omega_3)\}. \quad (3-8.8)$$

Thus one needs to compute

$$\begin{aligned} & \langle d\hat{x}(\omega_1) d\hat{x}(\omega_2) d\hat{x}(\omega_3) \rangle \\ &= \left\langle \overbrace{[d\hat{e}(\omega_1) + d\hat{\eta}(\omega_1) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega_1, \nu) d\hat{e}(\omega_1 - \nu) \hat{\eta}(\nu)]}^1 \right. \\ & \quad \overbrace{[d\hat{e}(\omega_2) + d\hat{\eta}(\omega_2) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega_2, \nu) d\hat{e}(\omega_2 - \nu)]}^2 \\ & \quad \left. \overbrace{[d\hat{e}(\omega_3) + d\hat{\eta}(\omega_3) + \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega_3, \nu) d\hat{e}(\omega_3 - \nu) \hat{\eta}(\nu)]}^3 \right\rangle \\ &= \langle 1Biii + 2Aiii + 3Aii + 3Bi + 1Cii + 2Ci \rangle \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\langle \hat{g}(\omega_3, \nu'') d\hat{e}(\omega_1) d\hat{e}(\omega_3 - \nu'') d\hat{\eta}(\omega_2) d\hat{\eta}(\nu'') \right. \\ & \quad + \hat{g}(\omega_3, \nu'') d\hat{e}(\omega_2) d\hat{e}(\omega_3 - \nu'') d\hat{\eta}(\omega_1) d\hat{\eta}(\nu'') \\ & \quad + \hat{g}(\omega_1, \nu) d\hat{e}(\omega_2) d\hat{e}(\omega_1 - \nu) d\hat{\eta}(\omega_3) d\hat{\eta}(\nu) \\ & \quad + \hat{g}(\omega_1, \nu) d\hat{e}(\omega_3) d\hat{e}(\omega_1 - \nu) d\hat{\eta}(\omega_2) d\hat{\eta}(\nu) \\ & \quad \left. + \hat{g}(\omega_2, \nu') d\hat{e}(\omega_1) d\hat{e}(\omega_2 - \nu') d\hat{\eta}(\omega_3) d\hat{\eta}(\nu') \right\rangle. \quad (3-8.9) \end{aligned}$$

There are six expectations above: the first one will be worked out fully. The remaining expectations involve nothing new.

$$\begin{aligned} & \left\langle \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega_3, \nu'') d\hat{e}(\omega_1) d\hat{e}(\omega_3 - \nu'') d\hat{\eta}(\omega_2) d\hat{\eta}(\nu'') \right\rangle \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega_3, \nu'') \langle d\hat{e}(\omega_1) d\hat{e}(\omega_3 - \nu'') \rangle \langle d\hat{\eta}(\omega_2) d\hat{\eta}(\nu'') \rangle \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{g}(\omega_3, \nu'') \langle d\hat{\epsilon}(\omega_1) d\hat{\epsilon}(\omega_3 - \nu'') \rangle \sigma^2 \delta'(\omega_2 + \nu'') d\omega_2 d\nu'' \\
&= \frac{\sigma^2}{2\pi} \hat{g}(\omega_3, -\omega_2) \langle d\hat{\epsilon}(\omega_1) d\hat{\epsilon}(\omega_3 + \omega_2) \rangle d\omega_2 \\
&= \frac{\sigma^4}{2\pi} \hat{g}(\omega_3, -\omega_2) \delta'(\omega_1 + \omega_2 + \omega_3) d\omega_1 d\omega_2 d\omega_3. \quad (3-8.10)
\end{aligned}$$

Here  $\omega_2$  has been chosen such that  $-\pi \leq \omega_2 \leq \pi$ . Due to the integration limits only the  $\delta'$  term shown in the fourth line above contributes. Further, the final (fifth) line is a consequence of  $\forall(x), d^2x = 0$ . Substitute this result (and the corresponding results for the remaining terms) into the complete equation to find

$$\begin{aligned}
&\langle d\hat{x}(\omega_1) d\hat{x}(\omega_2) d\hat{x}(\omega_3) \rangle = \\
&\sigma^4 / (2\pi) \{ \hat{g}(\omega_3, -\omega_2) + \hat{g}(\omega_3, -\omega_1) + \hat{g}(\omega_1, -\omega_3) \\
&\quad + \hat{g}(\omega_1, -\omega_2) + \hat{g}(\omega_2, -\omega_3) + \hat{g}(\omega_2, -\omega_1) \} \\
&\quad \delta'(\omega_1 + \omega_2 + \omega_3) d\omega_1 d\omega_2 d\omega_3. \quad (3-8.11)
\end{aligned}$$

Comparing this expression with MacDonald's equation (in radian frequency) one recovers the formula for the bispectrum in terms of the frequency coupling coefficients

$$\begin{aligned}
B(\omega_1, \omega_2) = \sigma^4 \{ &\hat{g}(\omega_3, -\omega_2) + \hat{g}(\omega_3, -\omega_1) + \hat{g}(\omega_1, -\omega_3) \\
&+ \hat{g}(\omega_1, -\omega_2) + \hat{g}(\omega_2, -\omega_3) + \hat{g}(\omega_2, -\omega_1) \}. \quad (3-8.12)
\end{aligned}$$

### 3-9 The Universal Bispectrum Model Might Arise in Practice, Almost.

The universal bispectrum model was written in Section 3-1 with little in the way of motivation. The unfortunate truth is that the model was developed on mathematical grounds and there is no known physically realistic system to which it might apply. This section shows how the obvious attempts to construct the universal bispectrum model come close but do not quite succeed.

One can write the general infinite-order moving average model as

$$x(t) = \epsilon(t) + \sum_{n=0}^{\infty} g(n) \epsilon(t - n), \quad (3-9.1)$$

just as in Section 3.1.

Here  $\epsilon(t)$  is a random variable (taken to be normally distributed, for simplicity) and values of  $\epsilon$  at different times are statistically independent.

Next assume that the weights  $g(n)$  themselves are stochastic functions of an independent time-dependent random variable  $\eta$ :

$$g(m) = \sum_{n=0}^{\infty} b_{nm} \eta(t - n - m). \quad (3-9.2)$$

Combining the above two equations gives

$$x(t) = \epsilon(t) + \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} b_{nm} \eta(t - n - m) \epsilon(t - n). \quad (3-9.3)$$

The above equation is almost the universal bispectrum model. There is a term  $\eta(t)$  missing. This flaw seems to afflict any effort to derive the universal bispectrum model in a natural fashion. "Derivation" of the model in a natural fashion is another opportunity for the reader!

It is possible to rewrite the model above in an equivalent auto-regressive form. If equation (3-9.1) is written as

$$x(t) = \epsilon(t) + \sum_{n=0}^{\infty} a(n) x(t - n) \quad (3-9.4)$$

with auto-regressive coefficients  $a(n)$  chosen to yield the same time series as given by the  $g(n)$  and if  $c_{nm}$  are chosen to relate to the  $a(n)$  as the  $b_{nm}$  relate to the  $g(n)$  then equation (3-9.3) becomes

$$x(t) = \epsilon(t) + \sum_{n=0}^{\infty} x(t - n) \sum_{m=0}^{\infty} c_{nm} \eta(t - n - m). \quad (3-9.5)$$

This model suffers the same defect as (3-9.3), but perhaps may suggest a defect-free model.

### 3-10 The Universal Bispectrum Model Can Be Clarified by a Diagram.

The primary concern of this chapter has been to determine how to compute a time series which corresponds to a given bispectrum. A physicist might call this the inverse problem or a statistician may call this the identification problem. A specific model which makes the solution of this problem possible (and easy) has been introduced and termed the universal bispectrum model. One method

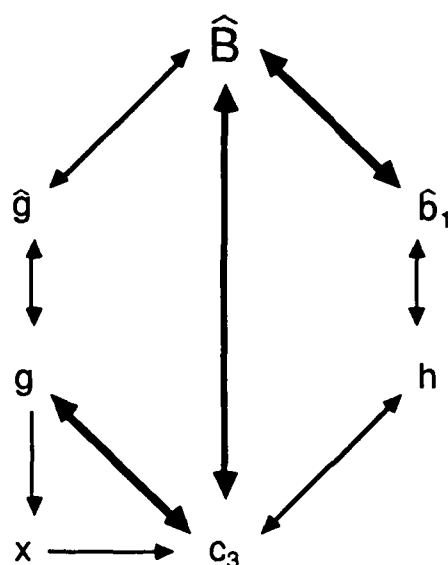


Figure 3-6 COMMUTATIVE DIAGRAM FOR UNIVERSAL BISPECTRUM MODEL.

of using this model is to desymmetrize the bispectrum to get its signature,  $\hat{b}_1$ , double Fourier transform the signature to get  $h$ , then resymmetrize to get  $c_3$ , and then desymmetrize to get  $g$ . The  $g$ 's then tell one how to compute a time series. A second method would reconstruct the  $\hat{g}$ 's from the original bispectrum, and then double Fourier transform to get the  $g$ 's and then get the time series. The various relationships and possible pathways are summarized in Figure 3-6.

Now it remains to see how this works in practice! Chapter 4 will present some necessary statistical tools and then Chapter 5 will present the model "under fire" .

## CHAPTER 4

### Estimation of the Bispectrum

#### 4-1 At Least Four Techniques are Used To Estimate the Bispectrum.

Now it is time to consider how the bispectrum can be estimated from sample data. Assume a discrete time series is somehow obtained. (If the model is inherently discrete, then the bispectrum is estimated for the entire principal domain. If the model is inherently continuous, then it should be anti-aliased filtered before sampling (that is, all frequencies higher than twice the sampled frequency should be removed) and it should be estimated only for the smaller "support set" as given in Chapter 2.) In either case, there are four techniques which are in common use for estimating the bispectrum.

The first technique, due to T. Subba Rao and M. Gabr [4-3] involves computing the triple correlation, smoothing it appropriately, and then transforming it to get the bispectrum. This is the natural approach based on the Section 1.2 definition of the bispectrum. The Gabr-Rao technique is documented in their book and will not be discussed further in this report.

The second technique is due to P. Huber and his co-authors [4-1]. Actually there are several variations in technique here, but the essential idea is to estimate the Fourier transform of the time series and then find the expectation of products of the transforms. This is the natural way to employ the definition of the bispectrum presented in Section 1.3.

The third technique is due to M. Hinich [4-8.1] and is closely related to that of P. Huber. The primary difference is that the averaging involved in estimating the Fourier transform and the subsequent averaging of these estimates to compute the bispectrum are distinct in the Huber approach and are performed together in the Hinich approach.

The fourth technique is very recent. It is due to M. Raghuveer [4-2] who claims it has much higher resolution than conventional methods (i.e., techniques 1-3 above) especially when used on short data sequences. The technique involves



assuming an underlying model for the process, estimating the parameters which characterize this model and then computing the bispectrum in terms of the estimated parameters. A somewhat odd feature of this technique is that two distinct models are used, the parameters for the model used for the bispectrum being *distinct* from those used to estimate the power spectrum. This technique will not be further discussed in this report, but Raghuveer's paper is recommended, not least for the bibliography which contains many recent applications of bispectral techniques.

The remainder of this chapter is devoted to the Hinich technique for estimating the bispectrum from sampled data. Particular attention is paid to both the theoretical underpinnings of the result (the author has great difficulty in using a formula without any idea of its origin) and its use in practice.

#### 4-2 One Must Distinguish the Sample Fourier Transform from the Quantity it Estimates – the Population Fourier Transform.

Before tackling the bispectrum, it is wise to consider a much simpler problem. In this post-FFT era, it is simple to *compute* the (discrete) Fourier transform of a given (discrete) time series. But what does this FFT tell one? If the sampled time series is deterministic, then the computed FFT is the true transform (aside from effects due to sampling interval and sample size). However, if the sampled time series is stochastic, then the computed FFT is just an *estimate* of the desired transform. If one were to take another sample one would get another estimate. Still it seems likely (and is true) that the expected value of the FFT estimate equals the true (or population) transform. What is surprising, however, is that the FFT is *not* a *consistent* estimator of the population transform. That is, no matter how long a time series one takes, the standard deviation of the FFT estimate for a given frequency does *not* approach zero (rather it remains constant). In other words, one cannot get increasingly good estimates of the population transform merely by increasing the sample size.

Estimating frequency domain quantities is therefore a more subtle matter than would appear at first sight. One must look at estimator *asymptotic bias* (does the expected value of the estimator approach the population quantity as sample size increases?) and *consistency* (does the estimator get better as one progressively increases the sample size – i.e., in the limit, does the variance of the estimator approach zero?). Moreover, one still has the concerns of the deterministic case: in particular, estimator *range* (the *Nyquist* frequency determines the maximum frequency for which an estimate can be made) and especially *resolution* (how closely spaced in frequency can one get reliable estimates?).

Ideally, one would like to have a procedure that, given tolerable bounds on bias and estimator standard deviation and specified resolution and frequency range, would specify the sampling parameters of the sampled time series (number of sampled series  $k$ , number of values in a single sample  $N$ , Nyquist frequency  $f_N$ , etc.). It may not be possible to physically attain the required sampling parameters – computational limits or sensor limitations may intervene. In this case, the procedure should provide some idea of what performance is possible.

The following sections derive an estimator which implements such a procedure – the Hinich estimator.

### 4-3 Fanfare for Brillinger's Formula.

This section is devoted to an exposition of one formula. This formula (if pressed) can answer virtually any question about the sampling properties of the FFT and is the key result underlying the Hinich estimator.

Suppose  $x(t)$  is a discrete time, zero mean, stationary stochastic time series. The discrete Fourier transform (DFT) of  $x$ , denoted  $\hat{x}$ , is defined by

$$\hat{x}(\omega_j) = \sum_{n=0}^{N-1} x(n) \exp[-i\omega_j n]. \quad (4-3.1)$$

The  $\hat{x}$  is a *sample* quantity, an estimate of the true or *population* spectrum. Each  $\hat{x}(\omega_j)$  estimates the Fourier amplitude at  $\omega_j = 2\pi j/N$ .

The statistical properties of the discrete Fourier transform estimate of the frequency spectrum are spelled out in a formula due to Brillinger and Rosenblatt. Brillinger presents this formula without emphasizing its usefulness. Here this formula gets its due.

The formula is actually an infinite set of equations, one equation for each number of sample quantities involved. On the left side of the formula is an expectation of products involving at most  $n$  sample quantities. On the right side is an expression involving the population quantities and the sample size  $N$ . It is possible to use this set of equations to determine the expectation of virtually any combination of DFT estimators (as will be seen shortly).

The left side expectation is the *cumulant* of  $n$  FFTs. Cumulants have been mentioned occasionally throughout this report, but only now is it time to take them seriously. Temporarily suffice it just to indicate a commonly used notation for cumulants:

$$[x_1, x_2, \dots, x_n]. \quad (4-3.2)$$

The next two sections provide the necessary background. Using this notation, Brillinger's formula [2-1] is

$$[\hat{x}_1, \dots, \hat{x}_n] = N\Delta(\omega_1 + \dots + \omega_n)f_n(\omega_1, \dots, \omega_{n-1}) + O(1). \quad (4-3.3)$$

Here  $\hat{x}_i$  denotes  $\hat{x}(\omega_i)$ ,  $N$  is the sequence length,  $n$  is the number of individual frequencies being looked at,  $\Delta(x)$  is defined as

$$\Delta(x) = \begin{cases} 1 & \text{if } x \bmod 2\pi = 0 \\ 0, & \text{otherwise} \end{cases}, \quad (4-3.4)$$

and  $f_n$  is the population cumulant spectrum defined by

$$f_n(\omega_1, \dots, \omega_{n-1}) = \sum_{u_i} [x(t + u_1), \dots, x(t + u_{n-1})] \exp[-i(\omega_1 u_1 + \dots + \omega_{n-1} u_{n-1})], \quad (4-3.5)$$

where the summation is over all integral values of the  $\{u_i\}$  and is independent of  $t$  by the assumption of stationarity.<sup>1</sup>

Equations (4-3.3) and (4-3.5) are simplified and less general versions of those presented in Brillinger. A factor of  $2\pi^{1-n}$  has been removed from equation (4-3.3) and this removes a corresponding factor from equation (4-3.5), which now provides the usual definitions of the population spectrum and bispectrum for  $n = 1$  and  $n = 2$ , respectively.

Basically Brillinger's formula states that, asymptotically (i.e., as  $N$  tends to infinity), the cumulant of the transform is the transform of the cumulant. This is not surprising: the bispectrum is initially defined as the transform of a cumulant (Section 1.2) but it is also equal to the cumulant of the transforms (Section 1.3). Thus one should expect that, asymptotically, the same relation will be true of the sampled quantities.

In the following sections, *the key to applying Brillinger's formula is to bear in mind the distinction between population and sampled quantities, and for the latter to pay attention to the highest power of  $N$  that appears.* Hopefully, this fact shall soon become clear.

(Note also that essentially the same simplified version of Brillinger's formula has been presented (and proved) in a readable paper by Kim and Powers. [4.3-1])

<sup>1</sup>The  $O(f(N))$  notation is standard and means that terms of order  $f(N)$  or lower may be present but have not been explicitly displayed. In this case, this notation means that constants (terms independent of  $N$ ) or terms that vary as an inverse power of  $N$  (i.e., terms which become smaller as  $N$  grows) are not explicitly shown. Nonetheless, for large  $N$ , the only important term is the one shown.

#### 4-4 A Diversion: Cumulants are Simply Expectations with Lower-Order Dependence Removed.

Treatment of cumulants has been postponed to this section because it is somewhat intricate and was not necessary until now. Further, the author hopes that the reader has become convinced of the importance of cumulants sufficiently to tolerate this section and the next. However, even here the precise general definition will not be given. Rather the remainder of this section will be pulled primarily from Doo Whan Choi's Ph.D. dissertation [4.4-1] which gives a particularly intuitive treatment of cumulants.

Cumulants are essentially expectations with lower order statistical dependence removed. If one wants the cumulant of  $n$  random variables, one constructs the expectation of the product of all  $n$  variables and then adds terms so that the net result will vanish if any subset of the variables is independent of any other subset.

For  $n = 2$ , this shows immediately that

$$[x_1, x_2] = \langle x_1 x_2 \rangle - \langle x_1 \rangle \langle x_2 \rangle, \quad (4-4.1)$$

since here one starts with the expectation  $\langle x_1 x_2 \rangle$  and it is clear that the right-hand side of equation (4-4.1) vanishes if  $x_1$  and  $x_2$  are independent.

For  $n = 3$ , one would then start with

$$[x_1, x_2, x_3] = \langle x_1 x_2 x_3 \rangle - \langle x_1 \rangle \langle x_2 \rangle \langle x_3 \rangle - \text{other terms} \quad (4-4.2)$$

which vanishes if  $x_1, x_2$ , and  $x_3$  are mutually independent. But suppose, following Choi, that only  $x_3$  is independent of  $x_1$  and  $x_2$  and that  $x_1$  and  $x_2$  are not independent. Then the right-hand side of equation (4-4.2) (neglecting the *other terms*) becomes

$$\langle x_1 x_2 \rangle \langle x_3 \rangle - \langle x_1 \rangle \langle x_2 \rangle \langle x_3 \rangle = [x_1, x_2] \langle x_3 \rangle. \quad (4-4.3)$$

Thus the term (4-4.3) must be included in the *other stuff* as well as the terms one gets by interchange of 3 with 1 and 3 with 2. The final definition of the cumulant of third order is

$$\begin{aligned} [x_1, x_2, x_3] = \langle x_1 x_2 x_3 \rangle &- \langle x_1 \rangle \langle x_2 \rangle \langle x_3 \rangle \\ &- [x_1, x_2] \langle x_3 \rangle \\ &- [x_1, x_3] \langle x_2 \rangle. \end{aligned} \quad (4-4.4)$$

One can see that this procedure may rapidly become cumbersome for larger  $n$  and that explicit rules are needed to efficiently compute these higher-order cumulants. However, these rules themselves are fairly complicated and the interested reader is referred to the references.

The remainder of this section is for the more knowledgeable or curious reader. Notice from the above equation that for the zero mean case the third-order moment and the third-order cumulant are the same. It is usually emphasized that the bispectrum is the Fourier transform of the *cumulant* rather than the *moment*, in spite of the fact that these quantities are equal.

If one views the *bispectrum* as the first in a sequence of higher order *polyspectra*, grounds for this distinction do emerge. In the higher order case, where cumulants and moments are distinct quantities, there are two reasons for choosing cumulants over moments.

First, cumulants have better independence properties than moments. Moments contain information about lower order moments, whereas cumulants are constructed in such a way that each order cumulant has the dependence on lower order ones removed. For example, it is meaningful to set all cumulants above the second order to zero (and thus get a Gaussian process), whereas it is impossible to set all moments above any specific order to zero. (It is true that these higher order moments contain no additional information and need not be specified for a Gaussian process, but they are *not* zero - that is just the point. Incidentally, there is a theorem which states that any non-Gaussian process has an *infinite* number of nonvanishing cumulants. Thus if the bispectrum of a particular process is non-zero, it must be true that some higher order polyspectra are also non-zero.)

Second, Brillinger has shown that for a commonly encountered class of processes, Fourier transforms of cumulants are better behaved than Fourier transforms of moments. He first shows that, if the moments and the cumulants are distinct, then either but not both of the corresponding transforms can be mathematically well-behaved (in the sense of being *proper* functions rather than requiring Dirac delta functions). Next he shows that for *ergodic* processes, the transforms of cumulants are proper functions and therefore transforms of moments must involve delta functions. (For an ergodic process the time average of any quantity for a single representative sequence equals the average of the same quantity at a fixed time but taken over all members of the ensemble. In effect, any representative sequence (or, *realization*) of an ergodic time series eventually explores the entire range of behavior typical of its ensemble. Moreover, ergodic processes are common both in theory - because it is convenient to be able to interchange ensemble and time averages - as well as in practice - because the "range of possible behavior" in typical ensembles is typically connected so that any one series *will* do anything possible for its entire class.)

#### 4-5 Everything We Need To Know: Table of Zero Mean Cumulants

There are many rather complicated formulae or sets of rules for evaluating cumulants of arbitrary order. For present purposes it is simpler to merely tabulate the cumulants that shall be used and refer the reader elsewhere for the general rules. The ones that shall be used in the following sections are those of orders less than or equal to six and for random variables with zero expectation. The notation used below is as follows. Rectangular brackets enclose cumulants; angular brackets enclose expectations. The braces with subscripted numbers are to be interpreted as instructions to replace the enclosed term with the sum of all distinct terms obtainable from the enclosed term by permutation of indices. The subscript outside the brace denotes how many terms should be obtained. (This notation is similar to that used by Dr. Choi in his dissertation, but it is not quite the same.)

$$\begin{aligned}
 [x_1] &= 0 \\
 [x_1, x_2] &= \langle x_1 x_2 \rangle \\
 [x_1, x_2, x_3] &= \langle x_1 x_2 x_3 \rangle \\
 [x_1, x_2, x_3, x_4] &= \langle x_1 x_2 x_3 x_4 \rangle - \{ \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle \}_3 \\
 [x_1, x_2, x_3, x_4, x_5] &= \langle x_1 x_2 x_3 x_4 x_5 \rangle - \{ \langle x_1 x_2 x_3 \rangle \langle x_4 x_5 \rangle \}_{10} \\
 [x_1, x_2, x_3, x_4, x_5, x_6] &= \langle x_1 x_2 x_3 x_4 x_5 x_6 \rangle - \{ \langle x_1 x_2 x_3 x_4 \rangle \langle x_5 x_6 \rangle \}_{15} \\
 &\quad - \{ \langle x_1 x_2 x_3 \rangle \langle x_4 x_5 x_6 \rangle \}_{10} \\
 &\quad + 2 \{ \langle x_1 x_2 \rangle \langle x_3 x_4 \rangle \langle x_5 x_6 \rangle \}_{15}
 \end{aligned}$$

#### 4-6 The Practice: Previous Results Allow One To Estimate the Power Spectrum.

The foregoing sections will now be put to good use. In this section, the statistics (expectation and variance) of the usual estimate for the power spectrum will be studied. In particular, the DFT will be shown to be an *inconsistent* estimator of the population spectrum. However, the main purpose of this section is to provide practice for the following section which applies the same techniques to the study of the statistical properties of the Hinich estimator for the bispectrum.

(Note that neither this section nor the following section actually carries through the analysis to the end, i.e., practical estimation of the spectrum and bispectrum. These sections are intended to provide only an understanding of the basic results underlying the practical techniques. Also note that all time series

shall be assumed to be zero mean so that the cumulants in the preceding table will be applicable.)

Define the "periodogram" estimator by

$$P_1(\omega_1) = \frac{\hat{x}(\omega_1)\hat{x}^*(\omega_1)}{N} = \frac{\hat{x}(\omega_1)\hat{x}(-\omega_1)}{N}. \quad (4-6.1)$$

It is more convenient to study the more general quantity

$$E_1(\omega_1, \omega_2) = \frac{\hat{x}(\omega_1)\hat{x}(\omega_2)}{N}. \quad (4-6.2)$$

The expectation of  $E_1$  is given by

$$\langle E_1(\omega_1, \omega_2) \rangle = \frac{\langle \hat{x}(\omega_1)\hat{x}(\omega_2) \rangle}{N} = \frac{[\hat{x}(\omega_1), \hat{x}(\omega_2)]}{N}, \quad (4-6.3)$$

where the second equality follows from the cumulant table.

By Brillinger's formula, this is

$$\langle E_1(\omega_1, \omega_2) \rangle = \Delta(\omega_1, \omega_2)f_2(\omega_1) + O(N^{-1}) \quad (4-6.4)$$

so that, asymptotically,

$$\langle P_1(\omega_1) \rangle = \langle E_1(\omega_1, -\omega_1) \rangle = f_2(\omega_1) \quad (4-6.5)$$

and examination of (4-3.3) shows that  $f_2$  is the population spectrum so that  $P_1$  is an unbiased estimator of the population spectrum. For future use, note that equation (4-6.3) indicates that

$$\omega_1 \neq \omega_2 \rightarrow \langle E_1(\omega_1, \omega_2) \rangle = O(N^{-1}), \quad (4-6.6)$$

and therefore estimates of amplitudes at different frequencies are asymptotically independent.

Next, the variance of the periodogram estimator is desired. One wants

$$\langle (P_1(\omega_1) - f_2(\omega_1))^2 \rangle = \langle (P_1(\omega_1))^2 \rangle - f_2^2(\omega_1). \quad (4-6.7)$$

The expectation of  $P_1^2$  is

$$P_2 = \langle (P_1(\omega_1))^2 \rangle = (1/N^2) \langle (\hat{x}(\omega_1)\hat{x}(\omega_1)\hat{x}(-\omega_1)\hat{x}(-\omega_1)) \rangle, \quad (4-6.8)$$

which is contained in the cumulant

$$c_4 = (1/N^2) [\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4], \quad (4-6.9)$$

where

$$\begin{aligned}\hat{x}_1 &= \hat{x}(\omega_1) \quad , \quad \hat{x}_2 = \hat{x}(\omega_1) \\ \hat{x}_3 &= \hat{x}(-\omega_1) \quad , \quad \hat{x}_4 = \hat{x}(-\omega_1).\end{aligned}\tag{4-6.10}$$

From the cumulant table

$$c_4 = P_2 - \{ \langle \hat{x}_1 \hat{x}_2 \rangle \langle \hat{x}_3 \hat{x}_4 \rangle \}_3, \tag{4-6.11}$$

where the  $1/N^2$  times the symmetrized sum in (4-6.11) is

$$\begin{aligned}& \langle \hat{x}(\omega_1) \hat{x}(\omega_1) \rangle \langle \hat{x}(-\omega_1) \hat{x}(-\omega_1) \rangle \\ & + \langle \hat{x}(\omega_1) \hat{x}(-\omega_1) \rangle \langle \hat{x}(\omega_1) \hat{x}(-\omega_1) \rangle \\ & + \langle \hat{x}(\omega_1) \hat{x}(-\omega_1) \rangle \langle \hat{x}(\omega_1) \hat{x}(-\omega_1) \rangle.\end{aligned}\tag{4-6.12}$$

Keeping only the highest order terms in  $N$  gives

$$(2/N^2) \langle \hat{x}(\omega_1) \hat{x}(-\omega_1) \rangle \langle \hat{x}(\omega_1) \hat{x}(-\omega_1) \rangle, \tag{4-6.13}$$

so that

$$c_4 = P_2 - (2/N^2) \langle \hat{x}(\omega_1) \hat{x}(-\omega_1) \rangle \langle \hat{x}(\omega_1) \hat{x}(-\omega_1) \rangle. \tag{4-6.14}$$

The  $\Delta$  function is automatically one since  $\omega_1 + \omega_2 + \omega_3 + \omega_4 = 0$ , so Brillinger's formula gives

$$c_4 = \frac{f_3(\omega_1, \omega_1, -\omega_1)}{N} + O(N^{-2}). \tag{4-6.15}$$

Therefore

$$\begin{aligned}\langle (P_1(\omega_1))^2 \rangle - f_2^2(\omega_1) &= (f_2(\omega_1))^2 \\ &+ (1/N) f_3(\omega_1, \omega_1, -\omega_1) + O(N^{-2})\end{aligned}\tag{4-6.16}$$

and, as  $N \rightarrow \infty$ ,

$$\langle (P_1(\omega_1))^2 \rangle - f_2^2(\omega_1) = f_2^2(\omega_1). \tag{4-6.17}$$

The right-hand side, being a population quantity, is independent of  $N$ , proving that the variance does not vanish asymptotically and that the "periodogram" estimator is not consistent. This result explains why smoothing and other techniques are required to estimate the spectrum.



A small diversion is now in order. Typically, the bispectrum is not presented *raw*; rather it is normalized. The normalized version is usually called the *bicoherence*. Powers defines the normalized bispectrum by

$$b^2(\omega_1, \omega_2) = \frac{|B(\omega_1, \omega_2)|^2}{\langle |\hat{x}(\omega)|^2 \rangle \langle |\hat{x}(\omega_1)\hat{x}(\omega_2)|^2 \rangle}. \quad (4-6.18)$$

This definition of bicoherence has the property that  $b$  is confined to range between 0 and 1 and is very nicely interpreted as the fraction of the total power at  $\omega$  which is due to the coupling. (This result is easy to derive: write down  $b^2/P$  and just substitute for these quantities using the formulae above.)

Power's definition of the bicoherence is not the only one possible: in fact, the bicoherence is more commonly defined by

$$b^2(\omega_1, \omega_2) = \frac{|B(\omega_1, \omega_2)|^2}{\langle |\hat{x}(\omega)|^2 \rangle \langle |\hat{x}(\omega_1)|^2 \rangle \langle |\hat{x}(\omega_2)|^2 \rangle}. \quad (4-6.19)$$

If the  $\hat{x}$ 's were population quantities rather than sample quantities, then the two definitions would be equivalent (the joint expectation in Power's definition would factorize as shown earlier). In general, there should not be much difference between the normalizations. However, the problem of choosing between them still remains. There is little discussion in the literature on this topic. The bulk of the *statistical* literature chooses the second form without much in the way of explanation. The author's preference is for Power's definition because it possesses the properties above.

#### 4-7 The Payoff: Previous Results Allow One To Estimate the Bispectrum.

The same type of argument can be used to derive the statistical properties of the Hinich estimator. This derivation in this section does not seem to have been published previously, however. Define the "raw bispectral" estimator by

$$F_1(\omega_1, \omega_2) = (1/N) \hat{x}(\omega_1)\hat{x}(\omega_2)\hat{x}(\omega_3), \quad (4-7.1)$$

where  $\omega_1 + \omega_2 + \omega_3 = 0$  and, for simplicity,  $\omega_1$  and  $\omega_2$  are taken to lie in the interior of the isosceles triangular subset of the principal domain. One now wants the expectation and variance of this estimator.

By analogy to Section 4.6, the expectation of  $F$  is obtained by relating it to a third order cumulant and then using Brillinger's formula to evaluate this

cumulant.

$$\begin{aligned}
 \langle F(\omega_1, \omega_2) \rangle &= \frac{\langle \hat{x}(\omega_1) \hat{x}(\omega_2) \hat{x}(\omega_3) \rangle}{N} \\
 &= \frac{[\hat{x}(\omega_1), \hat{x}(\omega_2), \hat{x}(\omega_3)]}{N} \\
 &= f_3(\omega_1, \omega_2) + O(N^{-1}), \quad (4-7.2)
 \end{aligned}$$

where the second equality follows from the cumulant table.

Next, the variance of the raw estimator is desired. One wants

$$\langle (F(\omega_1, \omega_2) - f_3(\omega_1, \omega_2))^2 \rangle = \langle (F(\omega_1, \omega_2))^2 \rangle - f_3^2(\omega_1, \omega_2). \quad (4-7.3)$$

The expectation of  $F^2$  is

$$\langle (F(\omega_1, \omega_2))^2 \rangle = (1/N^2) \langle (\hat{x}(\omega_1) \hat{x}(\omega_2) \hat{x}(\omega_3) \hat{x}(-\omega_1) \hat{x}(-\omega_2) \hat{x}(-\omega_3)) \rangle, \quad (4-7.4)$$

which is contained in the cumulant

$$c_6 = \frac{1}{N^2} [\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4, \hat{x}_5, \hat{x}_6], \quad (4-7.5)$$

where

$$\begin{aligned}
 \hat{x}_1 &= \hat{x}(\omega_1), & \hat{x}_4 &= \hat{x}(-\omega_1) \\
 \hat{x}_2 &= \hat{x}(\omega_2), & \hat{x}_5 &= \hat{x}(-\omega_2) \\
 \hat{x}_3 &= \hat{x}(\omega_3), & \hat{x}_6 &= \hat{x}(-\omega_3).
 \end{aligned} \quad (4-7.6)$$

From the cumulant table and Brillinger's formula,

$$\begin{aligned}
 c_6 &= (1/N^2) [a_1 + a_2 + a_3 + a_4] \\
 &= (1/N) f_6(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5) + O(N^{-2}), \quad (4-7.7)
 \end{aligned}$$

where

$$\begin{aligned}
 a_1 &= \langle \hat{x}_1 \hat{x}_2 \hat{x}_3 \hat{x}_4 \hat{x}_5 \hat{x}_6 \rangle = \langle |\hat{x}_1|^2 |\hat{x}_2|^2 |\hat{x}_3|^2 \rangle \\
 a_2 &= -\{ \langle \hat{x}_1 \hat{x}_2 \hat{x}_3 \hat{x}_4 \rangle \langle \hat{x}_5 \hat{x}_6 \rangle \}_{15} \\
 a_3 &= -\{ \langle \hat{x}_1 \hat{x}_2 \hat{x}_3 \rangle \langle \hat{x}_4 \hat{x}_5 \hat{x}_6 \rangle \}_{10} \\
 a_4 &= -\{ \langle \hat{x}_1 \hat{x}_2 \rangle \langle \hat{x}_3 \hat{x}_4 \rangle \langle \hat{x}_5 \hat{x}_6 \rangle \}_{15}. \quad (4-7.8)
 \end{aligned}$$

The above expressions can be simplified by keeping only the highest order terms in  $N$ . One knows that the highest order contribution comes from

those expectations with frequencies that sum to zero. Moreover the assumption that  $\omega_1$  and  $\omega_2$  lie inside the isosceles triangular subset of the fundamental domain implies that  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  are mutually unequal so that "coincidental" equalities are absent. One finds

$$\begin{aligned} a_1 &= \langle |\hat{x}_1|^2 |\hat{x}_2|^2 |\hat{x}_3|^2 \rangle \\ a_2 &= -[\langle |\hat{x}_1|^2 \rangle \langle |\hat{x}_2|^2 |\hat{x}_3|^2 \rangle + \langle |\hat{x}_2|^2 \rangle \langle |\hat{x}_1|^2 |\hat{x}_3|^2 \rangle \\ &\quad + \langle |\hat{x}_3|^2 \rangle \langle |\hat{x}_1|^2 |\hat{x}_2|^2 \rangle] \\ a_3 &= -|\langle \hat{x}_1 \hat{x}_2 \hat{x}_3 \rangle|^2 \\ a_4 &= 2\langle |\hat{x}_1|^2 \rangle \langle |\hat{x}_2|^2 \rangle \langle |\hat{x}_3|^2 \rangle. \end{aligned} \quad (4-7.9)$$

Examination of  $a_1$  and  $a_3$  validates the earlier claim that the expectation of  $F^2$  is contained within the cumulant. In fact, the entire desired variance in equation (4-7.3) is just  $(a_1 + a_3)/N^2$  so that combining equations (4-7.6), (4-7.7), and (4-7.9) gives

$$\langle (F(\omega_1, \omega_2))^2 \rangle - f_3^2(\omega_1, \omega_2) = -\frac{1}{N^2} [a_1 + a_4] + \frac{1}{N} f_6(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5) + O(N^{-2}). \quad (4-7.10)$$

The asymptotic behaviors of  $a_2$  and  $a_4$  are easily shown to be

$$\begin{aligned} a_2 &= 2N^3 f_2(\omega_1) f_2(\omega_2) f_2(\omega_3) + O(N^2) \\ a_4 &= -3N^3 f_2(\omega_1) f_2(\omega_2) f_2(\omega_3) + O(N^2), \end{aligned} \quad (4-7.11)$$

so that equation (4-7.10) simplifies to

$$\langle (F(\omega_1, \omega_2))^2 \rangle - f_3^2(\omega_1, \omega_2) = N f_2(\omega_1) f_2(\omega_2) f_2(\omega_3) + O(1). \quad (4-7.12)$$

This result is the key result which underlies the Hinich estimator and is equivalent to equation 2.5 of Hinich under the simplifying assumptions made here. (To get the complete result, it is only necessary to allow  $\omega_1$  and  $\omega_2$  to be arbitrary and then include the "coincidental" equalities which occur in evaluating the cumulants in (4-7.7).)

#### 4-8 The Mathematical Results Can Be Turned into a Practical Procedure for Estimating the Bispectrum.

The previous section has shown that  $F(\omega_1, \omega_2)$  is an unbiased but not consistent estimate of the bispectrum. It does not consistently estimate the bispectrum because its variance does not approach zero. In fact, by equation (4-7.12), the variance actually increases with  $N$ . In this section  $F(\omega_1, \omega_2)$  will be

turned into a practical (consistent) estimator. Unlike the earlier sections, the mathematics will be stated and not proved. This is due not to increased complexity, but to the lesser relevance of the mathematics for present purposes.

For simplicity of notation,  $F(\omega_j, \omega_k)$  will be denoted  $F(j, k)$ .

To obtain a consistent estimate, the  $F(j, k)$  is averaged over adjacent values in a square of  $M^2$  points centered at the points

$$\begin{aligned}(g_m, g_n) &= \left( \frac{(2m-1)M-1}{2}, \frac{(2n-1)M-1}{2} \right) \\ &= \text{where } m = 1, \dots, [N/(2M) + 0.5] \\ &= \text{and } n = 1, \dots, \min(m, [N/M - 2m + 3/2 + 3/(2M)])\end{aligned} \quad (4-8.1)$$

where  $M$  is determined by trading off resolution versus bias (as shall be seen below) and  $[ ]$  denotes the "greatest integer not exceeding" function.

The values of  $m$  and  $n$  above are constrained so that the center points are within the principal domain. The constraints on  $m$  and  $n$  are derived from the principal domain constraints on the set of  $(j, k)$ :

$$0 \leq j \leq N/2, 0 < k \leq j, 2j + k \leq N. \quad (4-8.2)$$

If not all points within the square are within the principal domain, only those points within the principal domain are included in the average. Thus the bispectrum estimator is given by

$$\hat{B}(m, n) = M^{-2} \sum_{j=(m-1)M}^{mM-1} \sum_{k=(n-1)M}^{nM-1} F(j, k), \quad (4-8.3)$$

subject to the constraints in equation (4-8.2).

If the bispectrum is slowly varying over the square, then this estimator is unbiased:

$$\langle \hat{B}(m, n) \rangle = B(f_{g_m}, f_{g_n}) + O(M/N). \quad (4-8.4)$$

If the power spectrum is slowly varying over the band of width  $M$  centered at the appropriate frequencies, then the variance of this estimator is given by

$$\text{Var } \hat{B}(m, n) = NM^{-4} Q(m, n) f_2(f_{g_m}) f_2(f_{g_n}) f_2(f_{g_m} + f_{g_n}) + O(M/N), \quad (4-8.5)$$

where  $Q(m, n) = M^2$  if the square is entirely within the principal domain; otherwise it is equal to the number of points  $(j, k)$  within the square but not on the

boundaries  $j = k$  or  $2j + k = N$  plus twice the number of points  $(j, k)$  on the boundaries.

From equation (4-8.5) it can be seen that the estimator given by equation (4-8.3) is a consistent estimator for values of  $M$  given by

$$\sqrt{N} < M < N. \quad (4-8.6)$$

The bias increases and the variance decreases as  $M$  increases. If the time series is divided into  $L$  segments, each of length  $N$ , and the bispectrum is estimated for each segment and all  $L$  estimates are averaged together, then if each of the  $L$  bispectrum estimates is uncorrelated, the variance of this coherent average is just

$$\text{Var } \hat{B}_L(m, n) = \text{Var } \hat{B}(m, n)/L. \quad (4-8.7)$$

In this case, consistency is obtained for values of  $M$  given by

$$\sqrt{N/L} < M < NL. \quad (4-8.8)$$

The asymptotic distribution of the estimator given by equation (4-8.3) is complex normal and independent for each frequency pair. Thus the distribution of the random variable

$$\chi^2(m, n) = \frac{2|\hat{B}(m, n)|^2}{\text{Var } \hat{B}(m, n)} \quad (4-8.9)$$

is noncentral chi-square with two degrees of freedom and noncentrality parameter

$$\lambda(m, n) = \frac{2}{NM^{-4}Q(m, n)}\gamma(f_{gm}, f_{gn}), \quad (4-8.10)$$

where

$$\gamma(f_{gm}, f_{gn}) = \frac{|\hat{B}(m, n)|^2}{f_2(f_{gm})f_2(f_{gn})f_2(f_{gm} + f_{gn})} \quad (4-8.11)$$

is called the skewness function. An estimate of the normalized bispectrum given in equation (4-8.9) can be obtained by using an estimate of the power spectrum in the expression for the variance. If the periodogram estimator for the power spectrum is used and it is smoothed over a band of at least  $M$  adjacent values, then this estimate of the normalized bispectrum will also be noncentral  $\chi^2$  distributed with two degrees of freedom and the same noncentrality parameter. If the bispectrum is coherently averaged over  $L$  segments, then the distribution of the normalized bispectrum after averaging is noncentral chi-square with 2 degrees of freedom and noncentrality parameter  $L\lambda(m, n)$ . If, on the other hand, the bispectrum

is normalized as in equation (4-6.19) and the normalized bispectrum is averaged over  $L$  segments, then the distribution of this incoherent average is noncentral chi-square with  $2L$  degrees of freedom and noncentrality parameter  $L\lambda(m,n)$ .

If the original time series is a linear process then it has constant skewness function. If it is a Gaussian process, the constant value is zero. Thus for a Gaussian process the asymptotic distribution of the estimator of the normalized bispectrum is a central chi-square with two degrees of freedom. This allows the values of the normalized bispectrum estimator to be thresholded and compared to values from a known distribution to determine when statistically significant values of the bispectrum estimator are occurring, i.e., values large enough that the assumptions of Gaussianity and or linearity can be rejected with some degree of confidence. Specific tests for Gaussianity and linearity have been developed using this estimator.

#### 4-9 The Practical Procedure Can Be Simplified To Give Order-of-Magnitude Estimates.

In practice, the formulae given in the previous section can be simplified even further. The estimator under consideration is obtained by starting with  $L$  sequences of length  $N$  each. For each sequence, one first constructs  $F(j,k)$  and then averages over squares with length  $M$  to get  $\hat{B}$ . The estimator is obtained by taking the average of  $\hat{B}$  over all  $L$  sequences.

One finds that the estimator bias is approximately

$$M/(NL) \quad (4-9.1)$$

and the estimator standard deviation is

$$\frac{1}{M} \sqrt{N/L} f_2^{3/2}. \quad (4-9.2)$$

#### 4-10 Estimating the Bispectrum Requires Smoothing.

In a nutshell, this chapter has shown that just as periodogram estimates need to be averaged to get a consistent power spectrum estimate, so must the raw bispectral estimates be similarly averaged. Moreover, a particular method for doing this averaging, the Hinich procedure, has been studied. The Hinich procedure averages adjacent raw bispectral estimates. It is also possible to average in the time domain to get a smoothed bicorrelation function and then transform. The results presented have been based on a formula by Brillinger which states that the transform of a cumulant is the cumulant of the transforms.

This chapter begins the treatment of practical issues which are needed to successfully apply the bispectrum. The following chapter extends this treatment to a worked example.

## CHAPTER 5

### The Model in Practice

#### 5-1 A Simple Example is Given by Choosing the Bispectrum Constant in a Specified Rectangular Region and Zero Elsewhere.

A simple example of a non-trivial bispectrum is one in which the bispectrum is zero except for a rectangular region where it is constant. The rectangular region is centered at  $\lambda_1, \lambda_2$  and its sides are parallel to the frequency axes with length  $2\Delta\lambda_1$  in the  $\omega_1$  direction and length  $2\Delta\lambda_2$  in the  $\omega_2$  direction. The magnitude of the bispectrum shall be denoted  $b$  and the bispectrum will be taken to have zero imaginary part. (See Figure 5.1.)

The complete bispectrum for this example, defined for the entire  $\omega_1, \omega_2$  plane, has numerous symmetries that have been presented in Chapter 2. These symmetries can be handled by the methods of Chapter 3. Here it suffices to compute the double Fourier transform of the function which looks exactly like that shown in Figure 5.1 and possesses none of the additional symmetries except for the  $2\pi$  periodicity in  $\omega_1$  and  $\omega_2$  and the conjugation symmetry.

This double Fourier transform is defined by

$$h(r, s) = \frac{1}{4\pi^2} \int \int \hat{h}(\omega_1, \omega_2) \exp[i(r\omega_1 + s\omega_2)] d\omega_1 d\omega_2, \quad (5-1.1)$$

where the integration limits are  $-\pi$  to  $\pi$  and

$$\hat{h}(\omega_1, \omega_2) = b \{ S(\omega_1; \lambda_1, \Delta\lambda_1) S(\omega_2; \lambda_2, \Delta\lambda_2) + S(\omega_1; \lambda_1, -\Delta\lambda_1) S(\omega_2; -\lambda_2, \Delta\lambda_2) \} \quad (5-1.2)$$

represents the rectangle function. (The second addend provides conjugation symmetry and is included to ensure reality of the double transform.) The function  $S$  is defined by

$$S(x; a; b) = \begin{cases} 1 & \text{for } a - b \leq x \leq a + b \\ 0 & \text{otherwise} \end{cases} \quad (5-1.3)$$



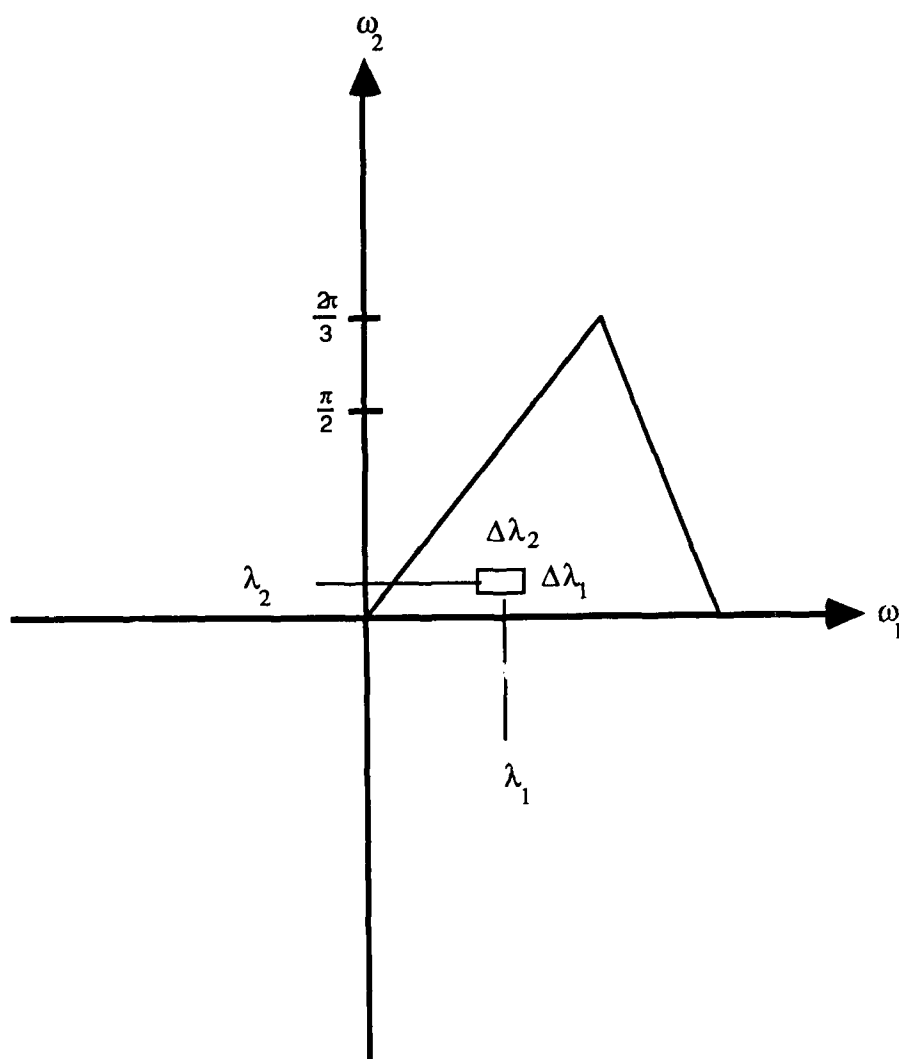


Figure 5-1 ARTIFICIAL BISPECTRUM EXAMPLE.

The Fourier transform in equation (5-1.1) is easy to evaluate. One has

$$h(r, s) = \frac{1}{4\pi^2} \left\{ \int_{\lambda_1 - \Delta\lambda_1}^{\lambda_1 + \Delta\lambda_1} \exp[ir\omega_1] d\omega_1 \int_{\lambda_2 - \Delta\lambda_2}^{\lambda_2 + \Delta\lambda_2} \exp[is\omega_2] d\omega_2 \right\} + c.c. , \quad (5-1.4)$$

where *c.c.* denotes the complex conjugate. Or, explicitly,

$$h(r, s) = \frac{2b}{\pi^2} \begin{cases} \Delta\lambda_1 \Delta\lambda_2 & ; r = s = 0 \\ \Delta\lambda_1 \cos(s\lambda_2) \sin(s\Delta\lambda_2)/s & ; r = s, s \neq 0 \\ \Delta\lambda_2 \cos(r\lambda_1) \sin(r\Delta\lambda_1)/r & ; r \neq 0, s = 0 \\ \cos(r\lambda_1 + s\lambda_2) \sin(r\Delta\lambda_1) \sin(s\Delta\lambda_2)/(rs) & ; r, s \neq 0 \end{cases} \quad (5-1.5)$$

If the rectangular region is restricted to lie entirely within the isosceles triangular subset of the fundamental domain, then the methods of Chapter 3 can be used to get the time domain correlation function  $c_3$ ,

$$c_3(r, s) = h(r, s) + h(s, r) + h(r-s, -s) + h(s-r, -r) + h(-s, r-s) + h(-r, s-r). \quad (5-1.6)$$

From this  $c_3$  one can determine the corresponding time domain coefficients  $g$  using equation (3-3.1) in reverse. Notice that  $g(r, s)$  will be non-zero for arbitrarily large  $r$  and  $s$ . In the computer implementation,  $g(r, s)$  will be truncated to zero past certain bounds  $R$  and  $S$ . This truncation will introduce a "ringing" effect, which is familiar from Fourier theory and which can be alleviated by using windowing techniques. However, that level of sophistication will not be needed here.

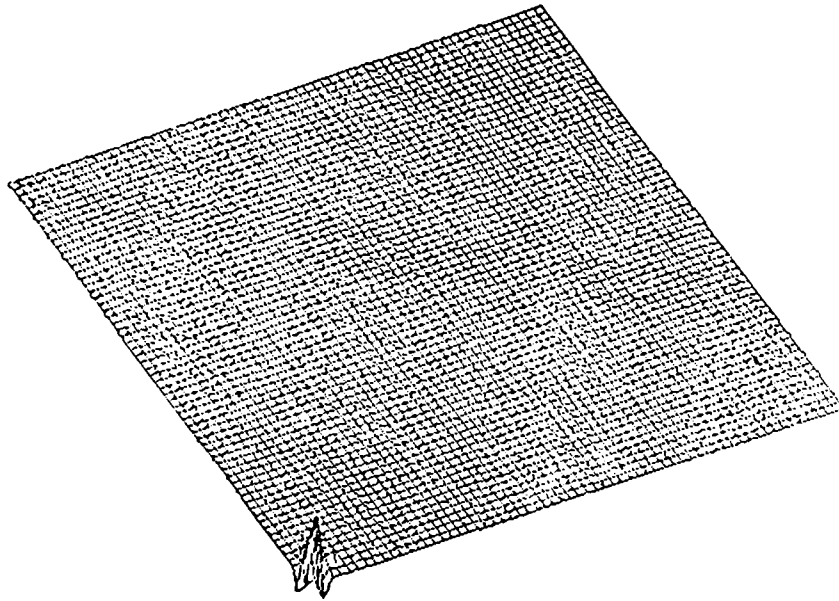
Thus one may produce the rectangularly shaped bispectrum with the discrete time series model below:

$$x(t) = \epsilon(t) + \eta(t) + \sum_{\substack{m=0 \\ n=0}}^{\infty} g(m, n) \epsilon(t-m) \eta(t-m-n), \quad (5-1.7)$$

with  $g$ 's computed using equation (5-1.6),  $h$ 's given by equation (5-1.5), and  $\eta(t)$  and  $\epsilon(t)$  independent zero-mean Gaussian random variables with standard deviation  $\sigma$ . Next, it is necessary to test the model.

## 5-2 Computer Simulations Give Good Agreement with the Model.

Testing the preceding model involves assembling a good deal of software. Moreover, any difficulties that may arise may be attributable either to inadequacies of the software or to mistakes in the mathematical development. To ensure



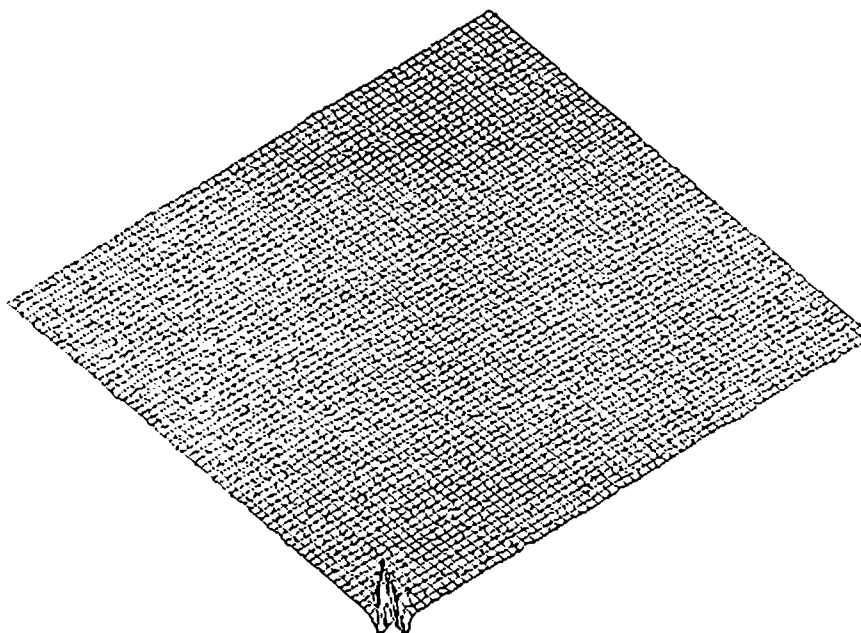
**Figure 5-2** THE "PREDICTED" BICORRELATIONS. (TEST MODEL)

that such was *not* the case, a very simple model was picked to start with. This model is defined by letting  $\sigma = 0.8$  and taking only a small number of  $g$ 's to be nonzero:

$$g = \begin{vmatrix} .4 & .7 & -.2 \\ .6 & -.5 & .3 \\ -.3 & .1 & .4 \end{vmatrix}. \quad (5-2.1)$$

These numbers were chosen more or less at random. Then 2000 time series of 128 points were generated from equation (3-1.2) using these parameter values. The  $g$  values are related very directly to the bicomrelations as shown in equation (3-3.4) so that the model's *predicted* bicomrelations are essentially an echo of the  $g$ 's. This is shown in Figure 5-2.

Given the model coefficients one can compute the predicted correlations using equation (3-2.2), the predicted power spectrum using equation (3-4.3), and the predicted bispectrum using equation (3-5.5). One may generate the "observed" quantities as follows. From the generated time series, compute the correlations and the bicomrelations directly from their definitions and then Fourier transform these quantities to get the corresponding power spectrum and bispectrum. (This method is not computationally efficient, but it has the advan-



**Figure 5-3** THE "OBSERVED" BICORRELATIONS. (TEST MODEL)

tages of simplicity and of reliability when developing software from scratch.) If everything is as it should be, the observed and the predicted quantities should be in close agreement.

These computations were done on a MicroVAX computer at Applied Research Laboratories (ARL:UT). The software which generates the time series and does the comparison of prediction versus observation is written in VAX Fortran and has been termed Bispectrum Workshop.

The results of Bispectrum Workshop are shown in the following set of figures. First, Figure 5-3 shows the observed bicorrelations. Next, side-by-side comparisons of the correlations, power spectra, and bispectra are shown in Figures 5-4, 5-5, and 5-6, respectively. The visual agreement is very good in all cases, and an examination of the actual numbers (not presented here, but available from Workshop output) confirms the match. Note also that the  $2\sigma^2$  term in equation (3-4.3) equals 1.28 which lies below, but near, the power spectrum curve as one might expect.

With this test successfully accomplished, it becomes desirable to try to reproduce a given bispectrum. A second program, ModelMaker, was written for this purpose. Rather than follow the analytical treatment in Section 5-1, this

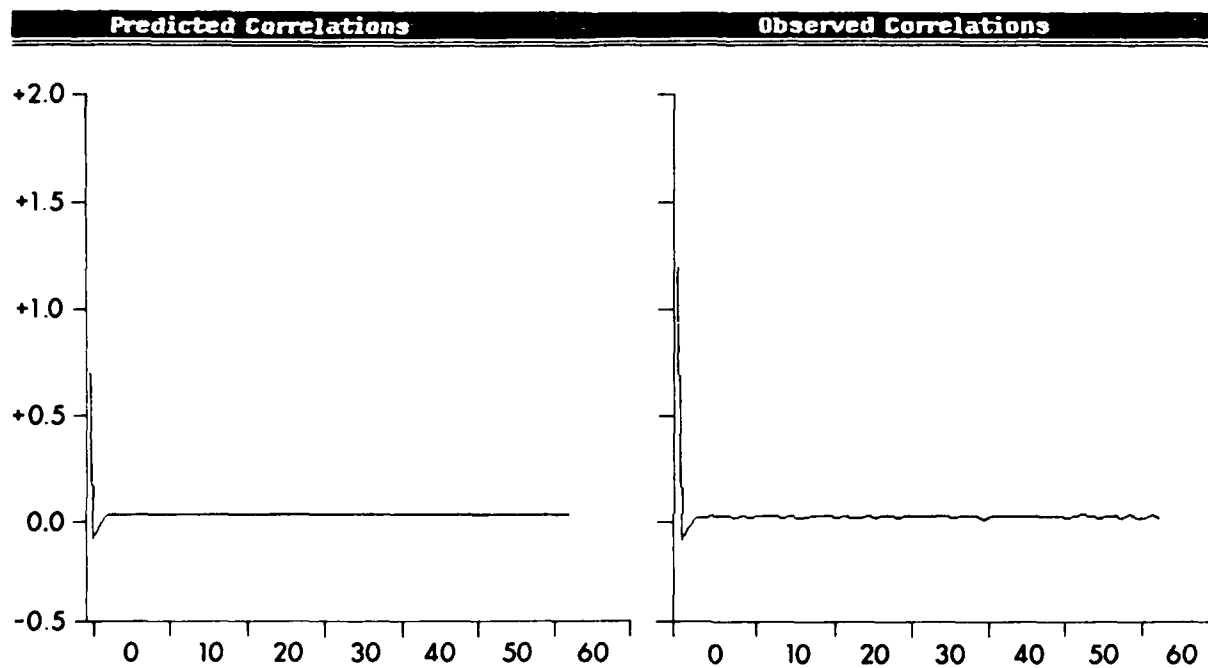


Figure 5-4 PREDICTED AND OBSERVED CORRELATIONS. (TEST MODEL)

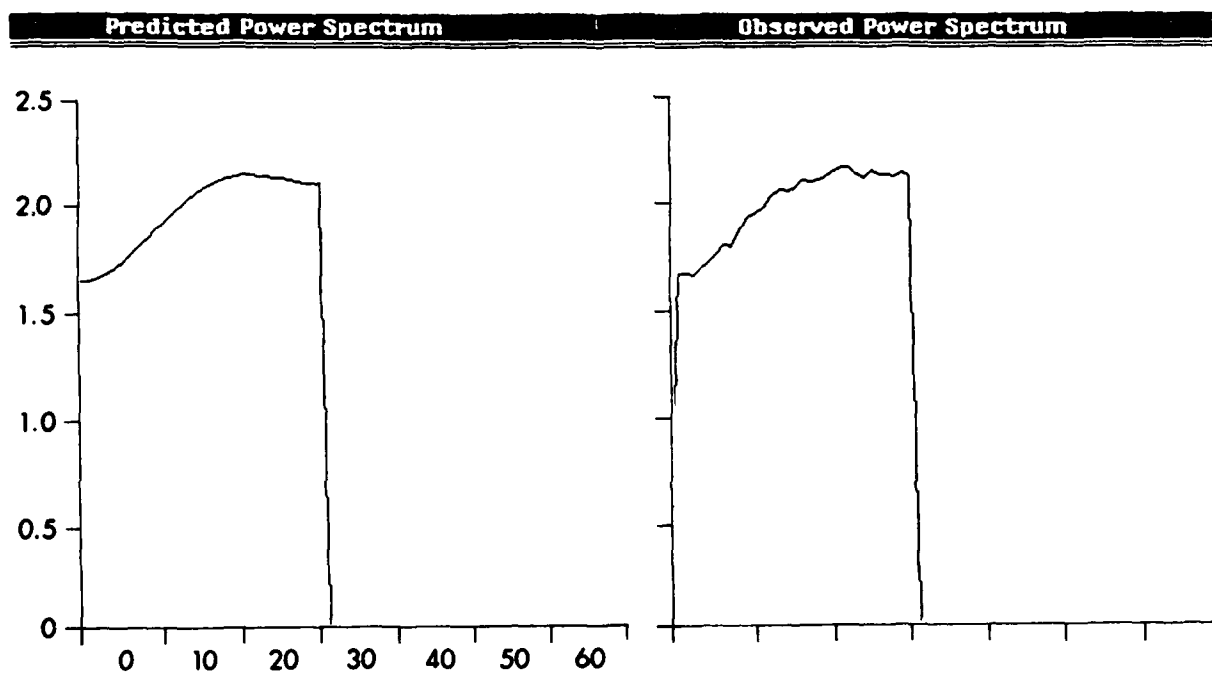


Figure 5-5 PREDICTED AND OBSERVED POWER SPECTRA. (TEST MODEL)

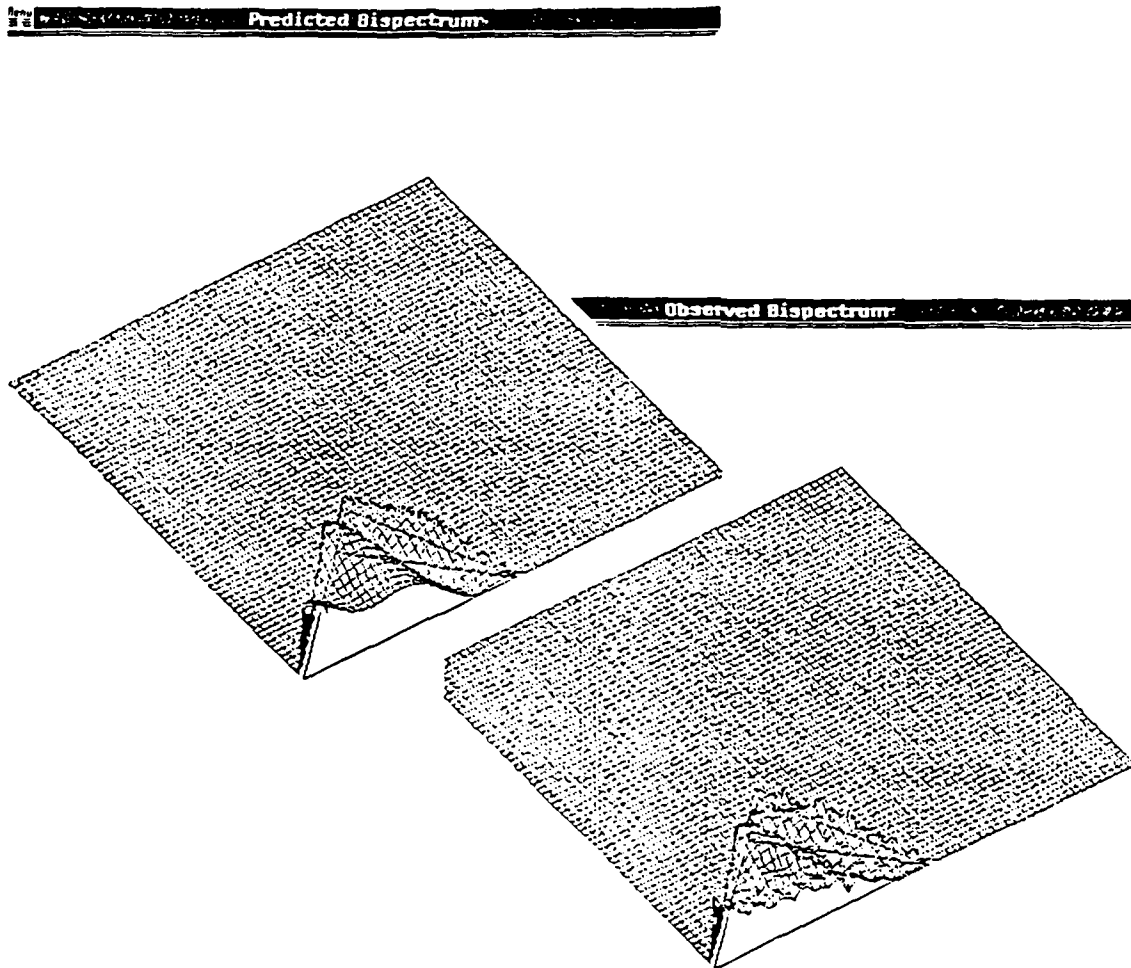


Figure 5-6 PREDICTED AND OBSERVED BISPECTRA. (TEST MODEL)

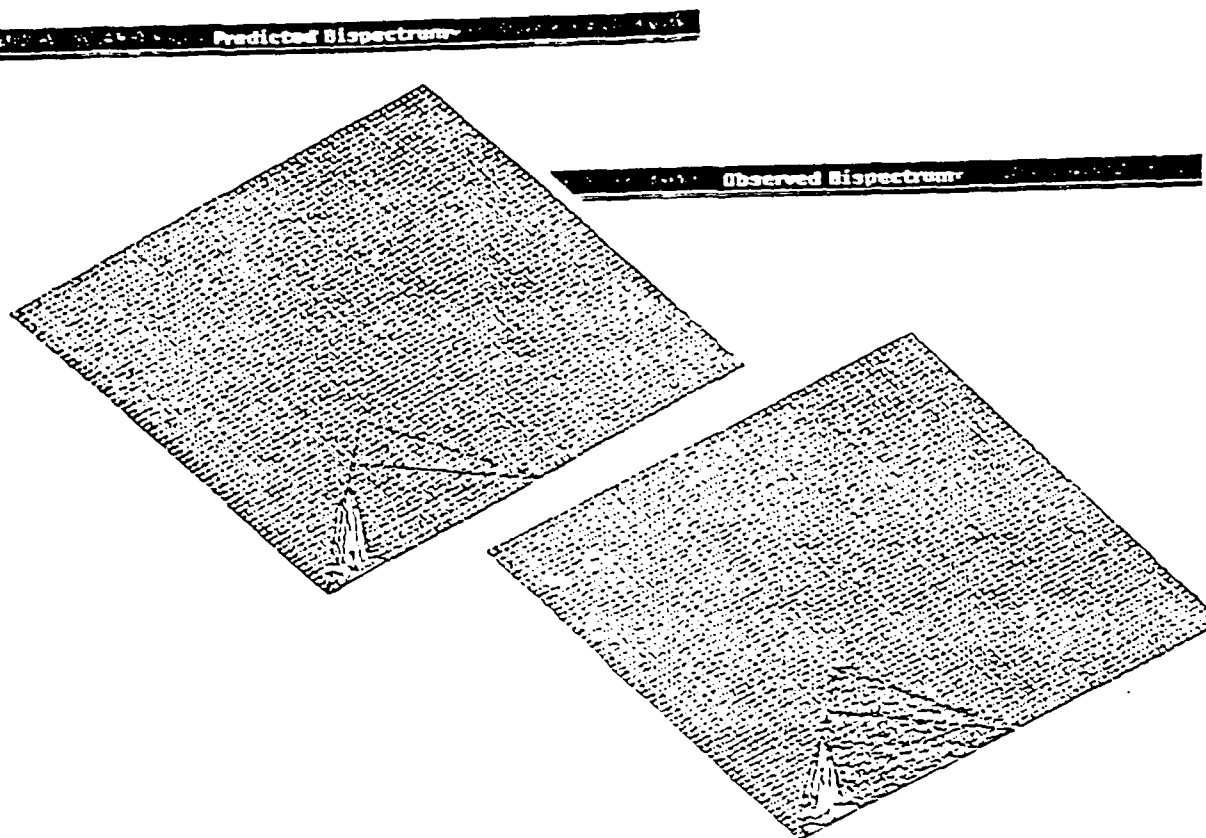


Figure 5-7 PREDICTED AND OBSERVED BISPECTRA. (SQUARE MODEL)

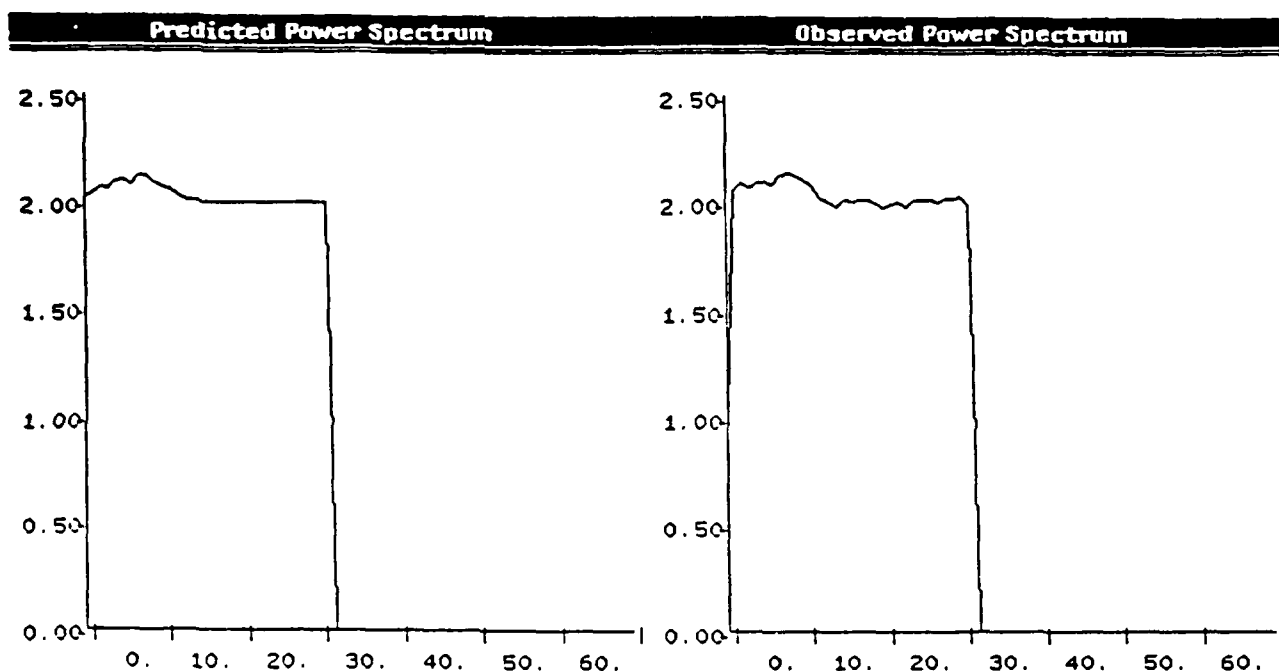
program takes a supplied continuous frequency bispectrum as input, samples it, double Fourier transforms to get the bicorrelation, and then computes the  $g$ 's from the bicorrelations. These  $g$ 's are output to a file which Workshop can read. (Note that the  $g$ 's are only determined if  $\sigma$  is assumed known. ModelMaker arbitrarily sets  $\sigma = 1$ . The following section will consider this matter in more detail.)

Here a model of the form shown in Figure 5-1 was used, with  $\lambda_1 = 3/16$ ,  $\lambda_2 = 9/32$ ,  $\Delta\lambda_1 = 5/128$ , and  $\Delta\lambda_2 = 15/128$ . In this case 5000 times series of 128 points were produced. However, the sampling was intentionally done very coarsely just to show sampling effects. Samples were taken at frequencies  $f_i = i\pi/32$ .

Side-by-side comparisons are shown in the following figures. First, notice that the predicted bispectrum is not flat; thus one cannot fault the observed bispectrum, which is in good agreement with it. (Figure 5-7.)

The non-flatness of the bispectrum is shared by the power spectra. Here, however, the limiting  $2\sigma^2$  value is 2 and this value is actually approached quite





**Figure 5-8** PREDICTED AND OBSERVED POWER SPECTRA. (SQUARE MODEL)

closely. (Figure 5-8.)

Finally, the predicted bicorrelations look much as one would expect for this model, and the observed bicorrelations can be seen to agree well. (Figure 5-9.)

Thus, computer simulation verifies the mathematics presented here and shows that this model can be used to invert the bispectrum. This section concludes with a few remarks concerning parameter values.

### 5-3 Reasonable Parameter Values are Obtainable through a Combination of Mathematical, Statistical, and Practical Considerations.

It is assumed that one wants to reproduce a given bispectrum with a model of the form given in equation (3-1.2). It is further assumed that one will

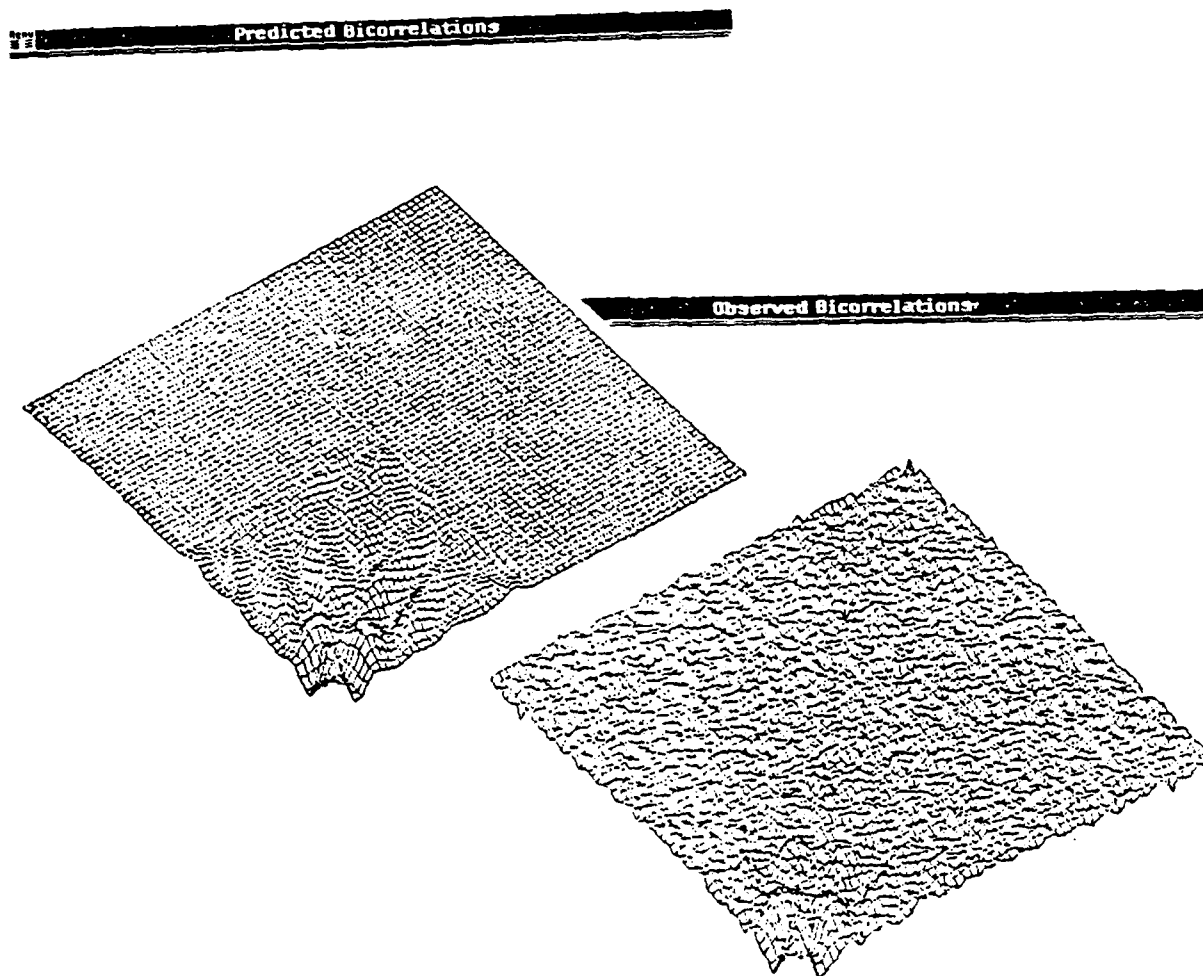


Figure 5-9 PREDICTED AND OBSERVED BICORRELATIONS. (SQUARE MODEL)

estimate the model bispectrum using the Hinich procedure described in Section 4. (The Hinich procedure is not used in the current implementation of Workshop as described in the preceding section.) The given bispectrum determines the  $g$ 's up to a scale factor. (Were  $\sigma$  known, this constraint would determine the scale factor as well.)

The model and estimator parameters at one's disposal are

$$\sigma, g_{scale}, R, S, M, N, \text{ and } L. \quad (5-3.1)$$

First, notice that the given bispectrum can be double Fourier transformed to get bicorrelations. These bicorrelations should be examined to see at which lag they begin to decay. The values of  $R$  and of  $S$  should be chosen to exceed this decay lag. Then the value of  $N$  (the number of points in a time series) should be chosen large with respect to  $R + S$ .

It is clear that in the present formulation,  $\sigma$  and an overall scaling factor  $g_{scale}$  can be traded off against one another to give a given bispectrum. This ambiguity can be employed to one's advantage by choosing the  $\sigma$  which yields the smallest estimator variance for a given bispectrum.

A time consuming, but exact, method for doing this would start by computing the variance of (4-8.5), using the expression for the power spectrum given in Chapter 3. This gives the variance at each point in the domain. One can now define some overall variance; e.g., average variance or perhaps variance weighted by the importance of regions of interest, to give a single number. Then one can vary  $\sigma$  and determine which value minimizes this quantity.

A simpler, but approximate, method would begin by approximating the magnitude of the given bispectrum by some number  $B$ . Next, write the  $g$ 's as

$$g(m, n) = g_{scale} g'(m, n), \quad (5-3.2)$$

where the  $g'$  are the  $g$  values which correspond to  $\sigma = 1$  (and hence are fixed for a given bispectrum). Now compute the quantity

$$\sum_{n=0}^{\infty} |\hat{g}'_n(\lambda)|^2 \quad (5-3.3)$$

and average it over  $\lambda$  to get the value  $G$ . With these approximations one can write the expression for the power spectrum as

$$f_2 = 2\sigma^2 + \sigma^4 g_{scale}^2 G. \quad (5-3.4)$$

Now  $\sigma$  and  $g_{scale}$  can be determined by minimizing  $f_2$  subject to the constraint that the bispectrum remain constant (or that  $\sigma^4 g_{scale} = B$ ). When one does this, one finds that

$$\sigma = [B^2 G]^{1/6}. \quad (5-3.5)$$

(And, for completeness, one gets  $g_{scale} = B^{7/3} G^{2/3}$ .)

The values of  $M$  and of  $L$  can now be put into the simplified estimator equations (4-9.1) and (4-9.2) to get adequately small bias and standard deviation.

#### 5-4 The Model Works!

This chapter demonstrates that the universal bispectrum model can be used to reproduce a desired bispectrum. The equations presented in the earlier chapters have been shown to be consistent with computer simulations. Some discussion of parameter values has been presented. The discussion as presented here is still incomplete, however.

There are alternate routes to actually employing the model which have not been discussed. The approach taken here is largely a time domain approach. A corresponding frequency domain approach may have some advantages. This frequency domain approach would derive the  $\hat{g}$  from the given bispectrum as before, but then instead of going directly to the  $g$ 's, the series would be computed in the frequency domain using (3-7.14). If desired, this series could then be transformed to get the time series. This approach is likely to be faster: it involves only a one-dimensional integration rather than a two-dimensional summation, and it is likely to be more accurate: agreement between prediction and observation does not depend on Fourier-transforming random quantities.

Further, the model can be naturally extended to give some degree of independent control of the power spectrum while retaining the capability of reproducing an arbitrary bispectrum. One obvious way of doing this is to introduce a third independent random Gaussian time series  $\zeta(n)$  and add a moving average contribution from this variable to the model (3-1.2). The new terms will not affect the bispectrum but will contribute additively to the power spectrum. Moreover this additional contribution to the power spectrum can be chosen at will. (One is stuck with at least as much power at each frequency as given by the un-adorned model, but one can add power as desired to re-shape the power spectrum without changing the bispectrum.) An alternative means of gaining this control is to add  $\eta(m)\epsilon(n)$  terms for  $m > n$ . These terms were neglected in the current formulation because they spoil the identifiability of the model and they allow no additional control of the bispectrum:  $\eta(m)\epsilon(n)$  and  $\eta(n)\epsilon(m)$  contribute additively to the

bispectrum so that one term can do duty for both. But these two symmetric terms do not contribute additively to the power spectrum. Therefore varying the summands while keeping the sum constant will change the power spectrum but not the bispectrum.

There are other interesting questions that should be explored. For example, is there a minimum power spectrum model for a given bispectrum, and, if so, is this model of the form of the universal bispectrum model? Can this approach be extended to higher order spectra? What does the continuous time formulation of this model look like? (In fact, the continuous time version will be a stochastic integral rather than a stochastic summation.) The author will concede that the the discussion of practical considerations given here needs to be amplified considerably.

## CHAPTER 6

### Summary

This report has attempted to provide an intuitive, uniform treatment of the bispectrum. Numerous perspectives on the bispectrum were presented with emphasis on physical and mathematical intuition. A limited catalog of examples is presented in an appendix to the report. A universal model, capable of generating any possible bispectrum, was worked out in both the time and frequency domains. An example of its use was presented. Attention was given to the symmetries, relationship between continuous and discrete time bispectra, and the mathematics underlying the sampling properties.

## APPENDIX A

### Catalog of Examples

### A-1 Linear Non-Gaussian White Noise Passed through Linear Device

This model and the next two are taken from the paper by Huber, Kleiner, Gasser, and Dumermuth [4-1]. Assume  $\langle X_i \rangle = 0$ ,  $\langle X_i^2 \rangle = 1$ , and  $\langle X_i^3 \rangle = \beta$ .

Define

$$Y_i = \sum h(i - j)X_j. \quad (\text{A-1.1})$$

Then, if

$$h(t) = \int_0^1 \exp(2\pi i \lambda t) H(\lambda) d\lambda, \quad (\text{A-1.2})$$

one finds

$$f_2(\omega) = |H(\omega)|^2 \quad (\text{A-1.3})$$

$$b(\omega_1, \omega_2) = \beta H(\omega_1) H(\omega_2) H^*(\omega_1 + \omega_2). \quad (\text{A-1.4})$$

### A-2 Gaussian Noise Passed through Squarer.

Assume  $X_i$  is Gaussian with  $\langle X_i \rangle = 0$  and  $\langle X_i^2 \rangle = 1$  and spectral density  $g(\lambda)$ . Pass this noise through a squarer:

$$Z_i = X_i + \alpha X_i^2, \quad (\text{A-2.1})$$

where  $Z_i$  is the output of the squarer,  $\alpha$  is the amplitude of the squared part and is numerically small.

The spectrum of  $Z_i$  is

$$f_2(\lambda) = g(\lambda) + 2\alpha^2 \int g(\mu)g(\lambda - \mu)d\mu \quad (\text{A-2.2})$$

and the bispectrum of  $Z_i$  is

$$f_3(\lambda_1, \lambda_2) = 2\alpha\{g(\lambda_1)g(\lambda_2) + g(\lambda_1)g(\lambda_1 + \lambda_2) + g(\lambda_2)g(\lambda_1 + \lambda_2)\} + \alpha^3 \text{ terms and higher.} \quad (\text{A-2.3})$$

### A-3 Poisson Activity

Define

$$Y_i = \sum_k h(t - Tk), \quad (\text{A-3.1})$$

where the  $T_i$  are Poisson-distributed times with

$$\langle T_{i+1} - T_i \rangle = \mu. \quad (\text{A-3.2})$$



Then, if

$$h(t) = \int_0^1 \exp(2\pi i \lambda t) H(\lambda) d\lambda, \quad (\text{A-3.3})$$

one finds

$$f_2(\omega) = \frac{1}{\mu} |H(\omega)|^2 \quad (\text{A-3.4})$$

and

$$b(\omega_1, \omega_2) = \frac{1}{\mu} \beta H(\omega_1) H(\omega_2) H^*(\omega_1 + \omega_2). \quad (\text{A-3.5})$$

These are the same results as in Example 1 provided  $\mu = 1$ .

#### A-4 Interacting Sinusoids

Let

$$x(t) = S_a(t) + S_b(t) + 2S_a(t)S_b(t) + N(t), \quad (\text{A-4.1})$$

where

$$\begin{aligned} S_a(t) &= \cos(\omega_a t + \phi_a) \\ S_b(t) &= \cos(\omega_b t + \phi_b). \end{aligned} \quad (\text{A-4.2})$$

We find that

$$\hat{X} \text{ is peaked at } \omega_a, \omega_b, \omega_a + \omega_b, \omega_a - \omega_b \quad (\text{A-4.3})$$

and

$$\hat{B} \text{ is peaked at } (\omega_a, \omega_b), (\omega_a, \omega_a - \omega_b). \quad (\text{A-4.4})$$

The bispectrum corresponding to this example is shown in Figure A-1.

#### A-5 Generalization of A-4, due to Powers.

Let  $y(t)$  be a time series and  $Y$  be its Fourier transform.

The power spectrum of  $y$ , as usual, is

$$P(\omega) = \langle Y(\omega) Y^*(\omega) \rangle \quad (\text{A-5.1})$$

and the bispectrum is

$$B(\omega_1, \omega_2) = \langle Y(\omega_1) Y(\omega_2) Y^*(\omega_1 + \omega_2) \rangle. \quad (\text{A-5.2})$$

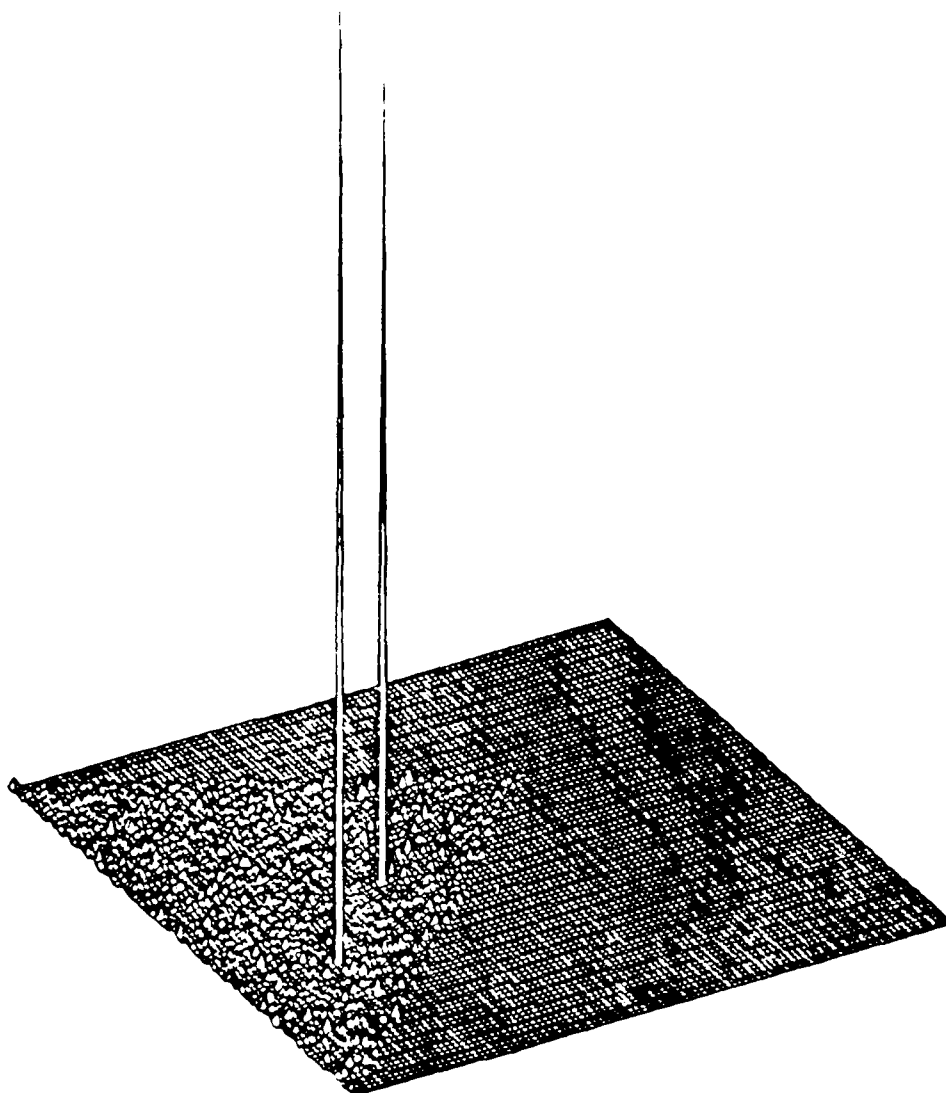


Figure A-1 INTERACTING SINUSOIDS.

This model [1.5-2] has nonlinearity of the simple form

$$Y(\omega) = AY(\omega_1)Y(\omega_2) + Y'. \quad (\text{A-5.3})$$

Here  $Y'$  contributes to  $Y(\omega)$  but is statistically independent of the coupling term.  $A$  represents the coupling constant.

Computing the power spectrum at  $\omega$ , one gets

$$P(\omega) = |A|^2 \langle |Y(\omega_1)Y(\omega_2)|^2 \rangle + \langle |Y'|^2 \rangle. \quad (\text{A-5.4})$$

The bispectrum at  $\omega_1, \omega_2$  is

$$B(\omega_1, \omega_2) = A^* \langle |Y(\omega_1)Y(\omega_2)|^2 \rangle. \quad (\text{A-5.5})$$

#### A-6 A Realistic Evolution Equation, due to Powers.

Here a realistic evolution equation is assumed:

$$\frac{\partial \phi_\omega}{\partial x} = V \phi_{\omega_1} \phi_{\omega_2} \exp(i\delta kx) \quad (\text{A-6.1})$$

with frequency matching ( $\omega = \omega_1 + \omega_2$ ), possible mismatch in wavenumber (characterized by  $\delta k \equiv k_{\omega_1} + k_{\omega_2} - k_\omega$ ), and wave-wave coupling (with coupling coefficient  $V$ ) [1.5-1].

It will be assumed that the above equation is true for some fixed  $\omega$  and that for the wave at this frequency  $\omega$  there is one pair of frequencies ( $\omega_1$  and  $\omega_2$ ) that dominates the integral. In particular the frequency dependence of  $V$  can be omitted.

In terms of the quantities of experimental interest, the above equation becomes

$$\frac{\partial Y_\omega}{\partial x} = V Y_{\omega_1} Y_{\omega_2} + ik(\omega) Y(\omega). \quad (\text{A-6.2})$$

It is a simple exercise (left to the reader) to take the above equation and derive

$$\frac{\partial P(\omega)}{\partial x} = VB(\omega_1, \omega_2) + V^* B^*(\omega_1, \omega_2). \quad (\text{A-6.3})$$

In terms of strictly real quantities  $V_R, V_I, B_R$ , and  $B_I$ , where

$$2V \equiv V_R + iV_I, \text{ and } B \equiv B_R + iB_I. \quad (\text{A-6.4})$$

This gives the wonderful equation

$$\frac{\partial P(\omega)}{\partial x} = V_R B_R(\omega_1, \omega_2) + V_I B_I(\omega_1, \omega_2). \quad (\text{A-6.5})$$

### A-7 A Stochastic Model of Gear Noises

This model is due to Sato and co-authors [Appendix-1] and applied to diagnosing gear noises for signs of wear. The model as given is a continuous time model. One has

$$x(t) = \sum_{n=0}^{\infty} a_n \cos(2\pi n f_0 t + \phi_n + d_n) \quad (\text{A-7.1})$$

where  $a_n$  are amplitudes of the various harmonics,  $\phi_n$  are deterministic phases, and the  $d_n$  are random phases.

The power spectrum for this model is

$$P(f) = \sum_{n=-\infty, n \neq 0}^{\infty} \frac{1}{4} a_n^2 \delta(f - n f_0), \quad (\text{A-7.2})$$

where  $\delta$  denotes the Dirac delta function.

The bispectrum is

$$B(f_1, f_2) = \sum_{m, n=-\infty, m, n \neq 0}^{\infty} 1/8 \ a_m a_n a_{m+n} \exp[i(\phi_m + \phi_n + \phi_{m+n})] \\ \langle \exp[i(d_m + d_n + d_{m+n})] \rangle \delta(f_1 - m f_0) \delta(f_2 - n f_0), \quad (\text{A-7.3})$$

so that one sees that the magnitude of the various terms is controlled by the expectation of the random phases. In particular, if the random phases are absent, then the expectation has its maximum value of one and the bispectrum shows the peaks most clearly.

### A-8 Brillinger's Polyspectrum Example.

This is a continuous time example [Appendix-2] which comes near to being a generalization of the universal bispectrum model. One takes

$$X(t) = \int a(t-u) dW(u) + \int \int b(t-u, t-v) dW(u) dW(v) \quad (\text{A-8.1})$$

where  $W(t)$  is a Wiener process,  $a$  has Fourier transform  $A$ ,  $b$  is assumed symmetric in its arguments and has double Fourier transform  $B$ .

The bispectrum of  $X(t)$  is

$$2[ \ A(\omega_1)A(\omega_2)B(-\omega_1, -\omega_2) + A(\omega_2)A(\omega_3)B(-\omega_1, -\omega_3) \\ + \ A(\omega_3)A(\omega_1)B(-\omega_3, -\omega_1) ] \\ + 8 \int \text{AvgPerm} \{ B(\omega, \omega_1 - \omega) B(\omega_2 + \omega, -\omega) B(\omega - \omega_1, -\omega - \omega_2) \} d\omega, \quad (\text{A-8.2})$$

and the notation  $AvgPerm\{f(\omega_1, \omega_2, \omega_3)\}$  denotes taking the average of  $f$  evaluated for all permutations of  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$ .

## REFERENCES

- [1.1-1] Peter V. Tryon. "The Bispectrum and Higher-Order Spectra: A Bibliography," National Bureau of Standards TN-1036 (April 1981).
- [1.2-1] David R. Brillinger. "An Introduction to Polyspectra," *Annals of Mathematical Statistics* **36**, 1351-1374 (1965).
- [1.2-2] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Springer-Verlag, New York, 1983).
- [1.2-3] J. Marcinkiewicz. *Math. Z* **44**, 612 (1939).
- [1.4-1] C. Chatfield. *The Analysis of Time Series: An Introduction. Second Edition* (Chapman and Hall, London, 1980).
- [1.4-2] Henry Abarbanel, Russ Davis, Gordon MacDonald, and Walter Munk. "Bispectra," JASON: The MITRE Corporation. McLean, Virginia. JSR-83-204 (January 1985).
- [1.4-3] Gordon MacDonald. "The Bispectra of Atmospheric Pressure Records," *Proceedings IBM Scientific Computing Symposia on Statistics* (IBM, White Plains, New York, 1963).
- [1.5-1] Young C. Kim and E. J. Powers. "Digital Bispectral Analysis and Its Applications to Nonlinear Wave Interactions," *IEEE Trans. Plasma Science* **PS-7** (2), 120-131 (1979).
- [1.5-2] Y. C. Kim, J. M. Beall, E. J. Powers, and R.W. Miksad. "Bispectrum and Nonlinear Wave Coupling," *Phys. Fluids* **23** (2), 258-263 (1980).
- [1.6-1] G. K. Batchelor. *The Theory of Homogeneous Turbulence* (Cambridge University Press, Cambridge, U.K., 1953).

- [1.6-2] K. N. Helland, K. S. Lii, and M. Rosenblatt. "Bispectra and Energy Transfer in Grid-generated Turbulence" (Chapter 3 of *Developments in Statistics*, Vol. 2) (Academic Press, New York, 1979).
- [1.6-3] T. T. Yeh and C. W. van Atta. "Spectral Transfer of Scalar and Velocity Fields in Heated-Grid Turbulence," *J. Fluid Mech.* **58** (2), 233-261 (1973).
- [1.6-4] K. S. Lii, M. Rosenblatt, and C. van Atta. "Bispectral Measurements in Turbulence," *J. Fluid Mech.* **7** (1), 45-62 (1976).
- [2-1] David R Brillinger. *Time Series Data Analysis and Theory* (Holt, Rinehart and Winston, Inc., New York, 1975).
- [2-2] David R. Brillinger and Murray Rosenblatt. "Computation and Interpretation of k-th Order Spectra." *Spectral Analysis of Time Series* (B. Harris, editor) (John Wiley and Sons, New York, 1970).
- [2-3] M. B. Priestley. *Spectral Analysis and Time Series, Volumes 1 and 2* (Academic Press, New York, 1981).
- [4-1] Peter J. Huber, B. Kleiner, B. T. Gasser, and G. Dumermuth. "Statistical Methods for Investigating Phase Relations in Stationary Stochastic Processes," *IEEE Trans. on Audio and Electroacoustics* **AU-19** (1), 78-86 (1971).
- [4-2] Mysore R. Raghuveer and Chrysostomos Nikias. "Bispectrum Estimation via AR Modeling." *Signal Processing* **10**, 35-48 (1986).
- [4-3] T. Subba Rao and M. M. Gabr. *An Introduction to Bispectral Analysis and Bilinear Time Series Models* (Lecture Notes in Statistics, Volume 24) (Springer-Verlag, Berlin, 1984).
- [4.3-1] Young C. Kim and Edward Powers. "Digital Bispectral Analysis and Its Application to Nonlinear Wave Analysis," *IEEE Trans. on Plasma Science* **PS-7** (2), (1979).
- [4.4-1] Doo Whan Choi. "Engineering Applications of Higher-Order Spectra" Unpublished Ph.D. dissertation. The University of Texas at Austin (1984).

[4-8.1] Melvin J. Hinich. "Testing for Gaussianity and Linearity of a Stationary Time Series," *J. Time Series Analysis* **3** (3) (1982).

[Appendix-1] Sato, Sasaki, and Nakamura. "Bispectral Analysis of Gear Noise," *J. Acoust. Soc. Am.* **62** (2) (1977).

[Appendix-2] David R. Brillinger. "An Introduction to Polyspectra," *Annals Mathematical Statistics* **36** (1965).



8 February 1988

**DISTRIBUTION LIST FOR  
ARL-TR-88-7  
UNDER ARL:UT INDEPENDENT RESEARCH AND  
DEVELOPMENT PROGRAM**

Copy No.

	<b>Commander</b>
	<b>Naval Air Systems Command</b>
	<b>Department of the Navy</b>
	<b>Washington, D.C. 20361-9330</b>
1	<b>Attn: W. Parigian (Code 933A)</b>
2	<b>CDR T. Santry (PMA 264-D)</b>
	 <b>Commander</b>
	<b>Naval Sea Systems Command</b>
	<b>Department of the Navy</b>
	<b>Washington, D.C. 20362-5101</b>
3	<b>Attn: Y. Yam (Code 63D4)</b>
4	<b>R. Lynch (Code 63D41)</b>
5	<b>M. Schultz (Code 63D22)</b>
6	<b>M. Montroll (PMS 394A)</b>
7	<b>C. Walston (Code 61Z137)</b>
8	<b>S. Wieczorek (Code 55NB)</b>
9	<b>E. Plummer (Code 63D1)</b>
	 <b>Commander</b>
	<b>Space and Naval Warfare Systems Command</b>
	<b>Department of the Navy</b>
	<b>Washington, D.C. 20363-5100</b>
10	<b>Attn: W. Richter (Code 80-B)</b>
11	<b>R. Cockerill (Code 180-521)</b>
12	<b>J. Bertrand (Code 183-1)</b>
13	<b>L. Parish (Code 180-4)</b>
14	<b>R. Mitnick (Code 180-5)</b>
15	<b>CAPT J. Jensen (Code 141)</b>
	 <b>Chief of Naval Research</b>
	<b>Department of the Navy</b>
	<b>Arlington, VA 22217-5000</b>
16	<b>Attn: J. Smith (Code 1211)</b>
17	<b>N. Gerr (Code 1111SP)</b>
18	<b>R. Fitzgerald (Code 1125AO)</b>
19	<b>K. Lackie (Code 132)</b>
20	<b>G. Remmers (Code 233)</b>
21	<b>J. Warfield (Code 234)</b>

Distribution List for ARL-TR-88-7 under ARL:UT Independent Research and Development Program  
(cont'd)

Copy No.

22	Chief of Naval Operations
23	Department of the Navy
	Washington, D.C. 20350-2000
	Attn: LT D. David (OP 092V)
	T. Long (OP 092V)
24	Commanding Officer
25	Naval Ocean Systems Center
26	San Diego, CA 92152
27	Attn: C. Persons (Code 732)
	T. Albert (Code 632)
	J. Silva
	D. Hanna
28	Commanding Officer
29	David Taylor Research Center
30	Carderock Laboratory
31	Bethesda, MD 20084-5000
32	Attn: B. Douglas (Code 1961)
33	J. Valentine (Code 1932)
	W. Hart (Code 1936)
	N. Keech (Code 1926)
	W. Reader (Code 1905.2)
	J. Niemiec (Code 1965)
34	Commanding Officer
	David Taylor Research Center
	Annapolis Laboratory
	Annapolis, MD 21402-5067
	Attn: L. Riddle (Code 1927)
35	Commanding Officer
36	Naval Intelligence Support Center
	4301 Suitland Rd.
	Washington, DC 20390
	Attn: J. Conrad (Code 20)
	C. Clark (Code 20)

Distribution List for ARL-TR-88-7 under ARL:UT Independent Research and Development Program  
(cont'd)

37  
38  
39

Commanding Officer  
Naval Air Development Center  
Warminster, PA 18974  
Attn: B. Steinberg (Code 5031)  
L. Allen (Code 5033)  
T. Posimonti

40  
41  
42  
43

Director  
Naval Research Laboratory  
Washington, DC 20375  
Attn: D. Bradley (Code 200)  
W. Kuperman (Code 5160)  
O. Diachok (Code 5120)  
P. Mignerey

44  
45  
46  
47  
48  
49  
50  
51  
52  
53

Commanding Officer  
Naval Underwater Systems Center  
New London Laboratory  
New London, CT 06320  
Attn: B. Cole (Code 33A)  
P. Herstein (Code 33A3)  
G. Mayer (Code 33501)  
R. Tozier (Code 2111)  
J. Martins (Code 2111)  
M. Maguire (Code 2111)  
W. Carey  
R. Tompkins (Code 33A2)  
R. Radlinski (Code 311)  
R. Dwyer

54  
55

Commanding Officer  
Naval Surface Warfare Center  
White Oak Laboratory  
Silver Spring, MD 20910  
Attn: E. Hein  
M. Williams

56

Commanding Officer  
Naval Oceanographic Office  
NSTL Station, MS 39522-5001  
Attn: W. Jobst (Code 7300)

57

Commanding Officer  
Naval Ocean Research and Development Activity  
NSTL Station, MS 39529-5004  
Attn: T. Goldsberry (Code 240)

Distribution List for ARL-TR-88-7 under ARL:UT Independent Research and Development Program  
(cont'd)

58	Defense Advanced Research Projects Agency
59	Department of the Navy
	Arlington, VA 22209
	Attn: C. Stuart
	K. Hawker
60-71	Commanding Officer and Director
	Defense Technical Information Center
	Cameron Station, Building 5
	5010 Duke Street
	Alexandria, VA 22314
	United States Naval Academy
72	Annapolis, MD 21402
	Attn: Library
	Director
	Applied Physics Laboratory
	Johns Hopkins University
	Johns Hopkins Rd.
73	Laurel, MD 20707
	Attn: F. Vaughan
	Department of Electrical Engineering
	The University of Texas at Austin
74	Austin, Texas 78712
	Attn: E. J. Powers
	Department of Aerospace Engineering
	The University of Texas at Austin
75	Austin, Texas 78712
	Attn: R. Stearman
	J. Bennighof
76	Signal Physics Group, ARL:UT
77	Garland R. Barnard, ARL:UT
78	Kevin W. Baugh, ARL:UT
79	W. Eugene Brown, ARL:UT

Distribution List for ARL-TR-88-7 under ARL:UT Independent Research and Development Program  
(cont'd)

80	George C. Coble, ARL:UT
81	Glen E. Ellis, ARL:UT
82	Marshall E. Frazer, ARL:UT
83	Thomas A. Griffy, ARL:UT
84	Melvin J. Hinich, ARL:UT
85	John M. Huckabay, ARL:UT
86	Clark S. Penrod, ARL:UT
87	F. Michael Pestorius, ARL:UT
88	Jack H. Sheehan, ARL:UT
89	Michael V. Smith, ARL:UT
90	Arnold J. Tucker, ARL:UT
91	Gary R. Wilson, ARL:UT
92	Murray A. Wolinsky, ARL:UT
93	Library, ARL:UT